THE UNIVERSITY *of* EDINBURGH

Title	Immunodominance, clonal composition and TCR[Bēta] repertoire of the bovine CD8$^+$ T-cell response to Theileria parva
Author	Connelley, Timothy
Qualification	PhD
Year	2007

Immunodominance, clonal composition and TCRβ repertoire of the bovine $CD8^+$ T-cell response to *Theileria parva*.

Timothy Connelley

Submitted in fulfilment of the requirements of the degree of Doctor of Philosophy

The University of Edinburgh – 2006

Declaration

I declare that the work presented in this thesis is my own original work, except where specified in the acknowledgements, and it does not include work forming part of a thesis presented successfully for a degree in this or another university.

Timothy Connelley

Abstract

In view of the evidence that $CD8^+$ T-cells are involved in mediating immunity against *Theileria parva*, the antigens recognised by these cells are obvious candidates for inclusion in a subunit vaccine. Results from previous studies have inferred that the $CD8^+$ T-cell response to *T. parva* is focused on a limited number of immunodominant antigens that exhibit polymorphism between different parasite strains. This could pose a major challenge to the design of a broadly effective subunit vaccine. The recent identification of CTL target antigens has provided the opportunity to characterise immunodominance within the *T. parva* system.

The objective of this study was to quantitatively assess immunodominance in the $CD8^+$ T-cell response to *T. parva* and to characterise the clonal composition and TCRβ repertoires of the epitope-specific T-cell populations. The results from four animals presented in this study have demonstrated that the $CD8^+$ T-cell response restricted by two MHC class I haplotypes is reproducibly dominated by single polymorphic epitopes. Using a suite of molecular tools developed during this study it was determined that the T-cell populations specific for both these epitopes were polyclonal but dominated by a limited number of large clonal expansions and the TCRβ repertoires expressed by these populations was diverse.

During the course of this work several novel bovine TCRβ genes were identified. Further examination of TCRβ cDNA transcripts and the bovine genome assembly substantially expanded the known bovine TCRβ repertoire, which is now the largest characterised for any species. Notably several Vβ subfamilies, especially Vβ1 and 13 have undergone extensive duplication and contain large numbers of genes. By annotating the available genomic data it has been shown that the bovine TCRB locus has a highly conserved synteny with the human TCRB locus. Furthermore, this annotation has demonstrated that prodigious duplication of a cassette containing a Vβ1 and Vβ13 gene has contributed to the large membership of these two subfamilies and that there are three D-J-Cβ clusters in the bovine TCRB locus rather than the two seen in the other mammalian TCRB loci described.

Acknowledgements

I must first acknowledge the unlimited support and attention that Jenna has given over the time we have known each other and been together. Without her ability to appear sincerely interested in my work it would have been infinitely harder to complete this thesis. Similarly, to all my long suffering friends and family, especially those also endeavouring to enter into research, who have shown boundless willingness to indulge my incessant discussion of my work, I am eternally grateful.

Thanks must be given to the various members of the laboratory who have all contributed in a multitude of ways. Prof. Morrison has provided endless, and often long, discussion on the work undertaken and given encouragement at several crucial interludes. Dr. N. M. MacHugh has supplied many reagents, without which the work presented here could not have been achieved, whilst Dr. F. Katzer has had the misfortune to share an office space with me and often had to fulfil the role of unofficial work-related agony uncle. Alison Burrells and Daniel Ngugi have provided both technical expertise and perhaps more importantly much needed sanity with their cheerfulness and words of wisdom. Others in the CTVM have also offered much needed help and assistance including Dr. Kim Picozzi, Dr. Eric Fevre and Anne Morrison. Thanks should also be given to Dr A. Bennett (Edward Jenner Institute for Vaccine Research) and Lynne Richardson (University of Oxford) for their technical advice on CDR3β-HDA and sequencing respectively.

Contents

List of figures

List of tables

Abbreviations and Definitions

Abbreviations

AgR	antigen-specific receptor
AMV-RT	Avian Myeloblastoma virus reverse transcriptase
APC	antigen presenting cell
bp	base pair
BSA	Bovine serum albumin
°C	degrees Celsius
C	Constant gene segment
Cβ	T-cell receptor beta chain constant gene segment
CD	Cluster of differentiation
cDNA	complementary deoxyribonucleic acid
CDR	complementarity determining region
CDR3β	T-cell receptor beta chain complementarity determining region three
CDR3β-HDA	TCRβ chain CDR3 heteroduplex analysis
cfu	colony forming unit
CLCN1	Muscle chloride channel 1
CO_2	carbon dioxide
cpm	counts per minute
^{51}Cr	Chromium radioactive isotope 51
CTL	cytotoxic lymphocyte
CTLp	cytotoxic lymphocyte precursor
CTVM	Centre for Tropical Veterinary Medicine
D	Diversity gene segment
dATP	2'-deoxyadenosine triphosphate
Dβ	T-cell receptor beta chain diversity gene segment
dCTP	2'-deoxycytidine triphosphate
dGTP	2'-deoxyguanosine triphosphate
dNTP	mixture of dATP, dCTP, dGTP and dTTP
dTTP	2'-deoxythymidine triphosphate
DBH	Dopamine-β-hydroxylase
DDW	double distilled water
DMSO	Dimethylsulphoxide
DNA	deoxyribonucleic acid
Eβ	TCRB locus enhancer
ECF	East Coast Fever
ELL	efferent lymph lymphocytes
EDTA	Ethylene-diamine-tetra-acetic acid
EPHB6	Ephrin type-b receptor 6 precursor
ER	endoplasmic reticulum
ERAP1	endoplasmic reticulum aminopeptidase 1
eSAG	endogenous superantigens
FACS	fluorescence-activated cell sorting
FAM	6-carboxy-fluorescein

FISH	fluorescent in-situ hybridization
FITC	fluorescein isothiocyanate
FR	framework region
g	grams
x g	force of gravity
h	hours
HCl	Hydrochloric acid
HDA	heteroduplex analysis
HIV	Human immuno-deficiency virus
HLA	Human leukocyte antigen
HUGO	Human Genome Organisation
I & T	infection and treatment
IAH	Institute for Animal Health
IDRP	insertion-deletion related polymorphisms
IFN	interferon
Ig	immunoglobulin
IL	interleukin
ILRAD	International Laboratory for Research on Animal Diseases
ILRI	International Livestock Research Institute
IMGT	Immunogenetics Database
IPTG	Isopropyl β-D-1-thiogalactopyranoside
^{111}In	Indium radioactive isotope 111
J	Joining gene segment
Jβ	T-cell receptor beta chain joining gene segment
Kb	kilo-base pairs
L	litre
LB	Luria-Bertani
LINE	long interspersed nucleotide elements
M	molar
mA	mili-amps
mAb	mono-clonal antibody
Mb	mega-base pairs
mM	mili-molar
MHC	Major-histocompatability complex
MHCI	Major-histocompatability complex class I
min	minute
µg	micro-grams
ml	milli-litres
µl	micro-litres
Mls-1a	Minor Lymphocyte Stimulating antigen 1a
mRNA	messenger ribonucleic acid
MVA	Modified vaccinia virus Ankara
MYA	million years ago
ORF	open reading frame
NaI	Sodium Iodide
NCBI	National Centre for Biotechnology Information
ND-PAGE	non-denaturing polyacrylamide gel electrophoresis

ng	nano-grams
NK	Natural-killer
nm	nano-metre
PBMC	peripheral blood mono-nuclear cells
PBS	Phosphate Buffered Saline
PDβ	Dβ promoter
PCR	polymerase chain reaction
PEG	Polyethylene glycol
pMHC	peptide-Major-histocompatability complex composite ligand
pmol	pico-molar
RAG	recombination-activating genes
RBC	red blood cells
RE	restriction endonuclease
RFLP	restriction-fragment length polymorphism
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
RSS	recombination signal sequence
RT-PCR	reverse transcriptase-polymerase chain reaction
RT	reverse transcriptase
s	second
SCM	standard culture medium
SDS	Sodium dodecyl sulphate
SSC	Sodium saline citrate buffer
snPCR	semi-nested polymerase chain reaction
TA	*Theileria annulata*-infected cell
TAE	Tris-acetate/EDTA electrophoresis buffer
TAP	transporter associated with Antigen Processing
TBE	Tris-borate/EDTA electrophoresis buffer
TCR	T-cell receptor
TCRβ	T-cell receptor beta chain
TCRβC	T-cell receptor beta chain constant gene segment
TCRβD	T-cell receptor beta chain diversity gene segment
TCRβJ	T-cell receptor beta chain joining gene segment
TCRβV	T-cell receptor beta chain variable gene segment
Tdt	Terminal deoxynucleotidyl transferase
TEMED	N, N,N', N-Tetramethylethylene-diamine
Tp	*Theileria parva*
TpM	*Theileria parva* (Muguga)-infected lymphocyte
Tris	2-amino-2-2(hydroxymethyl)-1,3-propanediol
V	volts
V	Variable gene segment
Vβ	T-cell receptor beta chain variable gene segment
v/v	volume to volume
w/v	weight to volume
WC	Workshop cluster
WHO-IUIS	World Health Organsiation-International Union of Immunological

	Societies
XGal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

Standard base pair (single letter code) and amino acid (both single and three letter code) abbreviations have been used throughout this thesis.

Definitions

The following definitions related to *Theileria parva* materials used in this study are derived from Irvin *et al.* (1983)

Isolate - viable organisms isolated on a single occasion from a field sample.

Stabilate - a sample of organisms preserved alive (usually in replicate) on a single occasion. In the context of this work sporozoite stabilates have been used.

Strain - a population of homogenous organisms possessing a set of defined characteristics. In the context of this work this term is used to refer to a population of organisms that have identical genetic material. Unambiguous characterisation of a strain can only be assured if the population of organisms was initiated from a parasite clone.

1 Chapter 1 - General introduction

1.1 *Theileria parva*

The obligate intra-cellular protozoan parasite *Theileria parva* is a member of the class Sporozoa, sub-phylum Apicomplexa, which includes several pathogens of medical and veterinary importance such as *Plasmodium* and *Toxoplasma.* It is the cause of an acute and usually fatal lymphoproliferative disease of cattle called East Coast Fever (ECF), which is prevalent in a large area of eastern, southern and central Africa (Figure 1.1). The predominant vector for *T. parva* transmission is the three host ixodid tick *Rhipicephalus appendiculatis* (Brown ear tick), although in its southern range the related species *R. zambeziensis* (Lawrence *et al.*, 1983) and *R. duttoni* (Lessard *et al.*, 1990) also serve as vectors.

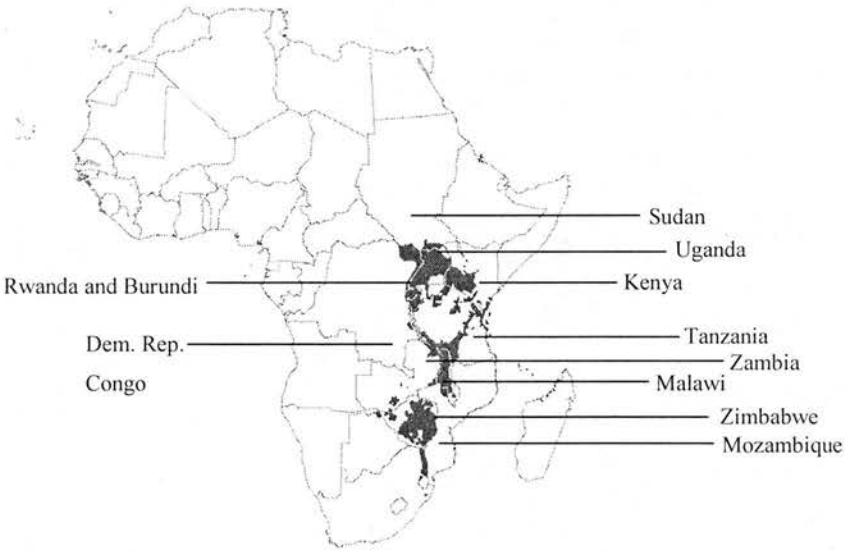


Figure 1.1 Geographical distribution of *Theileria parva* in Africa. Reproduced from the Annual Report of the International Laboratory of Research on Animal Diseases (ILRAD) 1990.

The disease is widely recognised as one of the major constraints on the livestock industries in the affected region. Firstly because of the financial cost of the disease and the measures taken to control it, estimated to be approximately US$ 168 million per year (Mukhebi *et al.*, 1992) and secondly because the more productive *Bos*

taurus breeds (european breeds such as Friesian) and their crosses with the indigenous *Bos indicus* breeds (e.g. Zebu) are more susceptible to ECF, and so increasing production through genetic improvement of the local livestock has been impeded.

1.1.1 Lifecycle of *T. parva* and pathogenesis of ECF

The lifecycle of *T. parva* is illustrated in Figure 1.2. Ticks become infected following ingestion of the intra-erythrocytic stage (piroplasms) whilst feeding on an infected animal. Within the tick gut, piroplasms rapidly differentiate (probably during the final stages of repletion) into micro- and macro-gametes which undergo syngamy to form spherical zygotes that invade the epithelial cells of the gut wall (Schein *et al.*, 1977). During the brief diploid phase of the lifecycle following syngamy there is an opportunity for sexual recombination (McKeever, 2006), which can facilitate the exchange within heterogenous *T. parva* populations of genes containing immunologically important determinants (section 1.1.3.3). Development from zygotes to motile kinetes and their subsequent migration through the haemocoel to the salivary glands appears to be synchronised with tick moulting (Young and Leitch, 1980, 1981), which may facilitate invasion of the salivary gland acinar cells by the kinetes (Mehlhorn and Schein, 1984). The kinetes have a strong tropism for the 'E' cells of type 3 acini (Fawcett *et al.*, 1982b; Young *et al.*, 1983), although other acini and cell types have been reported to contain kinetes (Mehlhorn and Schein, 1984; Shaw and Young, 1995). Within the salivary gland the kinetes transform into multinucleate, syncitial sporoblasts (Fawcett *et al.*, 1982a). Commencement of the next tick feed initiates sporogony, during which the sporoblasts rapidly develop and expand, until, after approximately 3-4 days of feeding, a terminal episode of cytoplasmic fission releases uni-nucleate sporozoites (Fawcett *et al.*, 1982a; Young *et al.*, 1983). Up to 50,000 sporozoites may be released from a single infected salivary gland (Fawcett *et al.*, 1982a), but this emission is gradual, so that inoculation of the sporozoites into the host occurs as a prolonged 'trickle' (Shaw and Young, 1995).

Following inoculation into the host, the parasite is not readily detectable for several days, thus events during this stage of the lifecycle have not been studied in detail *in vivo*. However, *in vitro* studies show that sporozoites rapidly enter (in as little as 3 minutes) into lymphocytes of all lineages – $CD4^+$, $CD8^+$ and γδ T-cells and B-cells (Baldwin *et al.*, 1988) by a process of receptor-mediated endocytosis (Fawcett *et al.*, 1982c; Webster *et al.*, 1985).

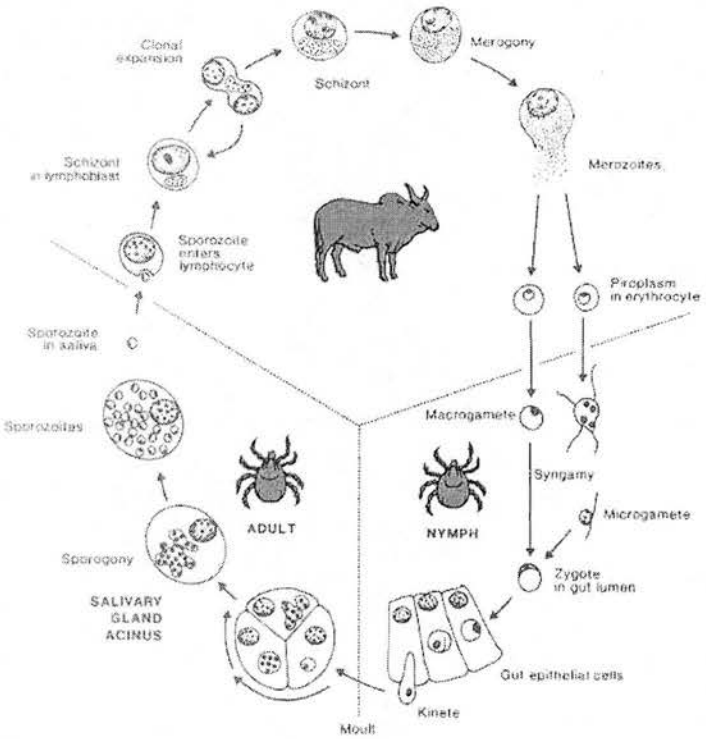


Figure 1.2 Lifecycle of *Theileria parva*. Reproduced from the Annual Report of the International Laboratory of Research on Animal Diseases (ILRAD) 1990.

The sporozoite rapidly penetrates the enclosing host-cell membrane, so that within 15 minutes it lies free within the cytoplasm, surrounded by a network of microtubules (Shaw, 2003; Shaw *et al.*, 1991). Within 3 days the sporozoite differentiates into a multi-nuclear schizont (Stagg *et al.*, 1981), a process that induces the infected cell to transform to a proliferating lymphoblast. The mechanisms

underlying this transformation are an active area of research and have been reviewed in several recent articles (Dessauge *et al.*, 2005; Dobbelaere and Kuenzi, 2004; Dobbelaere and Rottenberg, 2003). The schizont divides synchronously with the host cell (Hulliger *et al.*, 1964; Irvin *et al.*, 1982) and by close association with its mitotic spindle during division ensures schizont infection of both daughter cells and hence exponential clonal expansion of the parasite.

In experimental *in vivo* infections, schizonts are first detected in the draining lymph node 5-7 days after needle inoculation of sporozoites (Emery, 1981b; Morrison *et al.*, 1981; Shatry *et al.*, 1981), concurrent with the onset of pyrexia and localised lymphomegaly. Within 2-3 days schizonts disseminate throughout the lymphoid system and can be detected in distant lymph nodes (Emery, 1981b; Morrison *et al.*, 1981; Shatry *et al.*, 1981), causing a generalised lymphomegaly. The schizont-infected cells also metastasize in large numbers to non-lymphoid tissues including the gut and lungs and the subsequent tissue damage accounts for clinical symptoms of ECF such as diarrhoea and respiratory distress (a comprehensive clinical description of ECF is given in Irvin and Mwamachi, 1983). In the later stages of disease there is a severe leucopaenia (Morrison *et al.*, 1981; Shatry *et al.*, 1981), which is probably a consequence of both non-specific lympholysis (DeMartini and Moulton, 1973; Emery, 1981b; Emery *et al.*, 1981) and invasion of bone marrow tissue by schizont-infected cells (Morrison *et al.*, 1981; Wilde, 1966). Susceptible animals generally die within 3 weeks of infection, with death usually attributed to respiratory failure due to extensive pulmonary infiltration by parasitized cells and pulmonary oedema (Irvin and Mwamachi, 1983).

A proportion of schizonts eventually undergo merogony. The uni-nucleate merozoites produced are released by rupture of infected lymphocytes, invade erythrocytes and differentiate to piroplasms, which are usually evident from approximately day 12 post-infection (Shatry *et al.*, 1981). There is a limited but insignificant multiplication of piroplasms (Conrad *et al.*, 1986; Fawcett *et al.*, 1987) which constitute the infective stage for the tick vector.

1.1.2 Parasite Diversity

There exists a high level of genetic diversity within field populations of *T. parva*. In a recent study of 104 field isolates (Abbreviations and Definitions) of *T. parva* taken from 3 sites in Uganda, 84 distinct genotypes were identified (Oura *et al.*, 2005). Several *in vitro* methods have been developed to analyse genotypic polymorphisms among *T. parva* strains (Abbreviations and Definitions), including pulsed-field gel electrophoresis (Morzaria *et al.*, 1990), restriction fragment length polymorphism (RFLP) analysis using repetitive DNA, ribosomal or telomeric probes (Allsopp and Allsopp, 1988; Bishop *et al.*, 1993; Bishop *et al.*, 1996; Conrad *et al.*, 1987a), polymerase chain reaction (PCR) of Tpr1 repetitive DNA sequence (Allsopp *et al.*, 1989), PCR-RFLP (Bishop *et al.*, 2001) and most recently, a panel of micro- and mini-satellite markers (Oura *et al.*, 2003). Application of these techniques has permitted investigation of *T. parva* population structure and epidemiology in natural field situations (Bishop *et al.*, 1994; Oura *et al.*, 2005) and the epidemiological consequences of immunization interventions (Geysen *et al.*, 1999; Oura *et al.*, 2004). Currently, molecular genotyping is also being used to study sexual recombination between *T. parva* strains (F. Katzer, paper in process).

Evidence of *in vivo* immunological heterogeneity between *T. parva* strains was demonstrated over thirty years ago in cross-immunity trials of animals immunised by the infection and treatment (I & T) protocol (whereby animals are immunised by infection with a sporozoite stabilate (Abbreviations and Definitions) and concurrent treatment with oxytetracycline). Animals immunised using the I & T protocol were protected against homologous but not necessarily heterologous *T. parva*-strain challenge (Radley *et al.*, 1975c). Subsequently, *in vitro* recognition of *T. parva* schizont-infected cells by monoclonal antibodies (Conrad *et al.*, 1987b; Minami *et al.*, 1983) and $CD8^+$ (Goddeeris *et al.*, 1986; Goddeeris *et al.*, 1990) and $CD4^+$ (Brown *et al.*, 1989) T-cells has revealed antigenic polymorphism between strains. The existence of immunological heterogeneity has obvious implications for the future development of vaccines (section 1.1.5).

1.1.3 Protective Immunity to *T. parva*

Animals naturally recovered from infection with *T. parva* or immunized by the I & T protocol develop immunity to homologous challenge that persists for at least 3.5 years (Burridge *et al.*, 1972; Radley *et al.*, 1975a). The ability to generate protective immunity by immunizing animals with autologous schizont-infected cells (Emery *et al.*, 1982), and the transient appearance of schizont parasitosis in immune animals undergoing challenge (Morrison *et al.*, 1987) indicate that the protective immune response is directed against schizont-infected lymphocytes rather than sporozoites. The failure to transfer immunity with serum or γ-globulins from immunised animals (Muhammed *et al.*, 1975), or demonstrate *in vivo* or *in vitro* protective function of anti-schizont antibody (Creemers, 1982; Wagner *et al.*, 1974) suggests that protection is not mediated by antibody.

Demonstration that transfer of thoracic duct lymphocytes from immune to non-immune chimaeric twins was capable of transferring immunity (Emery, 1981a), identified the protective response as cell-mediated. *In vitro* and *ex vivo* studies identified a cell-mediated cytotoxicity directed against schizont-infected cells present in peripheral blood mononuclear cells (PBMC) of immune, but not naive, animals (Eugui and Emery, 1981; Pearson *et al.*, 1979). The appearance and activity of these cytotoxic cells in immunised animals undergoing challenge is chronologically linked to clearance of the parasite (Emery *et al.*, 1981; Eugui and Emery, 1981; Morrison *et al.*, 1987). The cytotoxic effector cells were characterised as being $CD8^+$ T-cells (Goddeeris *et al.*, 1986; Morrison *et al.*, 1987), and a subsequent study involving adoptive transfer of lymphocytes from immune to non-immune monozygotic twins undergoing challenge showed that transfer of $CD8^+$ T-cells was protective (McKeever *et al.*, 1994).

The frequency of cytotoxic T-cell precursors (CTLp) in immune animals can be very high. CTLp frequencies of 1:625 to 1:1005 PBMC can be found three weeks post-immunisation and levels of between 1:3084 and 1:3903 PBMC are maintained for a

period of up to 6 months (Taracha *et al.*, 1992). In immunised animals undergoing challenge, CTLp can constitute about 2.5% of the efferent lymph lymphocytes (ELL) draining from the regional lymph node by day 7 of challenge, and in the PBMC CTLp frequency peaks at approximately 0.1% on days 9-10 post-challenge (McKeever *et al.*, 1994).

1.1.3.1 *T. parva*-specific cytotoxic lymphocytes show parasite strain-specificity and major histocompatability class I restriction

The lack of absolute cross-protection offered by I & T immunization indicates that the protective immune response exhibits a degree of parasite strain-specificity (Radley *et al.*, 1975c). Study of *ex vivo* and *in vitro* responses have demonstrated that *T. parva*-specific cytotoxic lymphocytes (CTLs) display parasite strain-specificity (Goddeeris *et al.*, 1986; Goddeeris *et al.*, 1990; Morrison *et al.*, 1987; Taracha *et al.*, 1995a; Taracha *et al.*, 1995b). Results from a detailed *in vitro* analysis of CTL clones derived from *T. parva* (Muguga)-immunised animals indicate that CTLs are generally able to recognise some, but not all heterologous strains (Goddeeris *et al.*, 1990).

In a study of 22 animals that were immunized by the I & T protocol with either *T. parva* (Muguga) or *T. parva* (Marikebuni 3219) strains and then challenged with the reciprocal parasite, Taracha *et al.* (1995a) demonstrated a correlation between *T. parva*-strain specificity of CTLs *in vitro* and cross-immunity protection *in vivo*. Immunization in 9 of the animals generated cross-protection and these animals developed CTL capable of recognising both *T. parva* (Muguga) and *T. parva* (Marikebuni 3219). The remaining 13 animals did not generate cross-immunity, suffering severe disease following heterologous parasite strain challenge, and analysis revealed that their CTL recognised the immunizing but not the challenge parasite strain.

As would be anticipated for $CD8^+$ T-cells, *T. parva*-specific CTLs are subject to restriction by major histocompatability class I (MHCI - Morrison *et al.*, 1987). A notable feature of the MHCI restriction of *T. parva*-specific CTL responses in

MHCI heterozygous animals is that the response is usually preferentially biased towards one of the MHCI haplotypes (Morrison *et al.*, 1987; Taracha *et al.*, 1995b) and furthermore, in haplotypes expressing more than one MHCI product they are frequently biased towards a single MHCI gene product (Morrison *et al.*, 1987; Taracha *et al.*, 1995b).

1.1.3.2 Immunodominance

The strain-specificity of *T.parva*-specific CTL responses and the finding that they are often restricted by a single MHCI product implies that the $CD8^+$ T-cell response in individual animals is focused on a limited numbers of immunogenic epitopes that display polymorphism between strains.

Moreover, the presence of an apparent hierarchy in MHCI haplotypes, with respect to their capacity to preferentially restrict *T. parva*-specific CTL, suggests that these epitopes exist in an immunodominance hierarchy (section 1.3), whereby certain epitopes are capable of dominating the elicitation of T-cell responses (Morrison, 1996a; Morrison *et al.*, 1987).

Further evidence of immunodominance in the $CD8^+$ T-cell response to *T. parva* is provided by the results from a study of monozygotic twins homozygous for the bovine MHCI haplotype A10 (expressing the two MHCI products A10 and KN104). In this study one twin was immunised using *T. parva* (Muguga) and the other with *T. parva* (Marikebuni 3219). The *T. parva* (Marikebuni 3219)-immunised twin generated cross-reactive KN104-restricted CTLs and was protected from challenge with *T. parva* (Muguga), whereas the *T. parva* (Muguga)-immunised twin generated KN104 restricted *T. parva* (Muguga)-specific CTLs and suffered severe clinical reaction to subsequent *T. parva* (Marikebuni 3219) challenge (Taracha *et al.*, 1995b). However, following recovery this twin developed cross-reactive KN104-restricted CTL. These results indicate that in the context of KN104-restriction, although an epitope common to both strains is capable of eliciting a $CD8^+$ T-cell response, there is an immunodominant *T. parva* (Muguga)-specific epitope (Taracha *et al.*, 1995b),

the presence of which inhibits the CTL response against the common epitope (section 1.3.2.4).

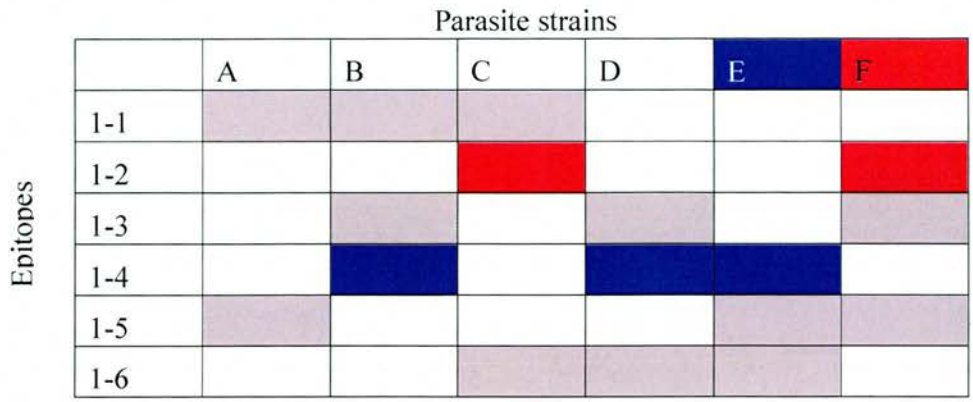
1.1.3.3 The mosaic model

In vitro analysis of the cross-reactivity of a panel of 31 CTL clones derived from *T. parva* (Muguga)-immunised animals of various MHCI haplotypes demonstrated that the strain-specificity of $CD8^+$ T-cells is profoundly influenced by their MHCI-restriction (Goddeeris *et al.*, 1990). Clones restricted by the same MHCI element displayed identical strain-specificity, whereas CTL of disparate MHCI-restriction exhibited different parasite-strain recognition patterns. Based on these data it has been proposed that CTL epitopes are distributed amongst the different *T .parva* strains according to a mosaic model (McKeever and Morrison, 1998; Morrison, 1996a).

A simplified form of the mosaic model in presented in Figure 1.3. Six parasite strains (A-F) each express 2 or 3 potentially immunogenic epitopes restricted by the MHCI haplotypes 1 (Figure 1.3a) and 2 (Figure1.3b). Each epitope is present in only some of the strains and the epitopes are more or less randomly distributed amongst the strains. This is consistent with the location of polymorphic epitopes in different parasite proteins, the genes of which have the potential to segregate during sexual recombination (section 1.1.1), which the model assumes is a frequent event between heterologous parasite strains co-present in infected ticks. As the epitopes presented by MHCI haplotype 1 (1-1 to 1-6) and haplotype 2 (2-1 to 2-6) may be located in different proteins their distribution amongst the strains differs. The model assumes that immunized animals generate CTL against only the most immunodominant epitope present in the immunising strain.

Applying the model allows the prediction of the cross-reactivity of the CTL induced by immunisation with each strain in the two MHCI haplotypes. For example, immunization with strain F will generate CTL capable of also recognising strain C in haplotype 1 animals and CTL capable of recognising strains A + D in haplotype 2 animals; whereas immunisation of haplotype 1 and 2 animals with strain E will

A - MHCI haplotype 1



B - MHCI haplotype 2

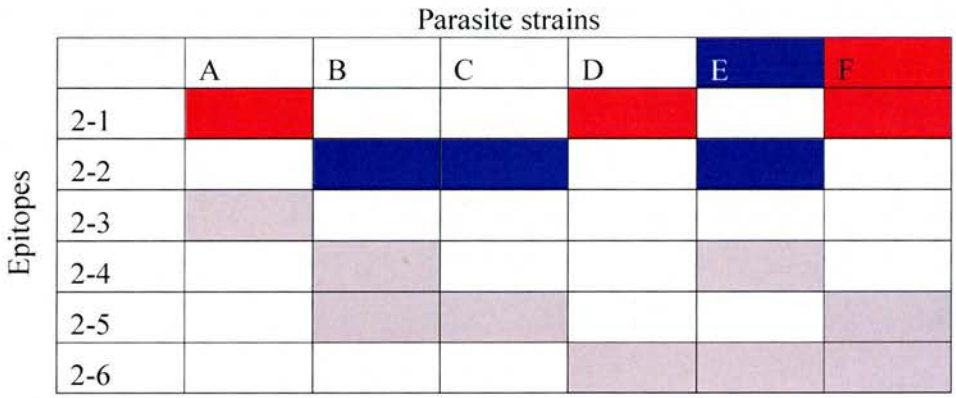


Figure 1.3 Schematic representation of the mosaic model. The epitopes expressed by each parasite strain (A-F) are represented by the coloured boxes. Figure 1.3a and Figure 1.3b show the distribution of epitopes restricted by MHCI haplotype 1 (1-1 to 1-6) and MHCI haplotype 2 (2-1 to 2-6) respectively. The CTL response following immunisation is assumed to be directed only to the most immunodominant epitope expressed in the immunizing strain with the immunodominance hierarchy of epitopes as follows; 1-1>2-1>1-2>2-2>1-3>2-3>1-4>2-4>1-5>2-5>1-6>2-6. Blue and red boxes indicate the epitopes against which CTL responses are directed following immunization with strains E and F respectively.

generate CTL that will cross-react with strains B + D and strains B + C respectively. Thus, in agreement with experimental observations, the model shows that both the immunising parasite strain and the MHCI phenotype of an animal will influence the

cross-reactivity of the resulting CTL response and therefore its ability to protect against heterologous challenge (Goddeeris *et al.*, 1990; Taracha *et al.*, 1995b). The model also facilitates consideration of MHCI heterozygous animals - in the illustrated example immunization of a MHCI 1/2 animal with strains E and F will induce CTL cross-reactive with strains B + C and A + D respectively - and replicates the dominance of certain MHCI haplotypes in restricting the CTL response to particular parasite strains (Morrison *et al.*, 1987; Taracha *et al.*, 1995b).

According to the mosaic model no strain is anticipated to generate inherently broad cross-reactive CTL. In agreement with observations from experimental (Dolan *et al.*, 1980; Morzaria *et al.*, 1987; Radley *et al.*, 1975b) and field data, the model predicts that certain combinations of two or more immunising parasite strains can provide substantially wider coverage - if it is assumed that the immunodominant epitope from each strain is capable of eliciting CTL as a consequence of presentation by different antigen-presenting cells (APCs) permitting subversion of immunodominance (section 1.3.2.4). For example, immunization with a combination of strains E and F would protect MHCI haplotype 2 animals against all heterologous strain challenge and haplotype 2 animals against challenge with all but strain A.

1.1.4 Other components of the immune response

The majority of published work has focused on studying the protective $CD8^+$ T-cell response to *T. parva*. However, some studies have provided insight into the responses of other components of the immune system.

A strong proliferative $CD4^+$ T-cell response is seen *in vitro* following stimulation of PBMC from immunised cattle with autologous *T. parva*-infected cells (Baldwin *et al.*, 1987). Although some *T. parva*-specific $CD4^+$ T-cells have cytolytic activity (Baldwin *et al.*, 1992) it is thought their main function is to provide help to the activation of *T. parva*-specific $CD8^+$ T-cells. Evidence from a study conducted by

Taracha *et al.* (1997) suggests that $CD4^+$ T-cell help is required for the activation of both naïve and memory *T. parva*-specific $CD8^+$ T-cells *in vitro*; whilst production of soluble factors (most likely to be cytokines such as IL-2) by $CD4^+$ T-cells appears sufficient to activate memory $CD8^+$ T-cells, naïve cells seem to require direct physical contact with parasite-specific $CD4^+$ T-cells. These results are consistent with the need for $CD4^+$ T-cell help seen in many $CD8^+$ T-cell responses (Bevan, 2004). Results from cell-transfer experiments with enriched *T. parva*-specific $CD4^+$ T-cells suggest that they are not able to confer protection in the absence of other effectors (unpublished observations reported in McKeever *et al.*, 1999).

It has also been established that γδ T-cell populations expand *in vivo* during the course of infection and *in vitro* when stimulated with autologous *T. parva*-infected cells; a proportion of these γδ T-cells show cytotoxic activity which is neither MHC-restricted nor strain-specific (Daubenberger *et al.*, 1999). There is some evidence that natural killer-like (NK-like) cells may also contribute to the immune response *to T. parva* (Goddeeris *et al.*, 1991); however the significance of the γδ T-cell and NK-like cell responses against *T. parva* have not been evaluated.

Although humoral responses do not appear to play a significant role in immunity induced by a single immunisation with *T. parva* (Muhammed *et al.*, 1975), animals that receive multiple challenges develop antibodies capable of neutralising the infectivity of sporozoites *in vitro* (Musoke *et al.*, 1982). *In vitro* studies have demonstrated the ability of these antibodies to neutralise sporozoites of different parasite strains (Dobbelaere *et al.*, 1984; Musoke *et al.*, 1984) and subsequent work has shown these antibodies recognise a major 67 KDa surface coat protein of sporozoites (p67) encoded by a single copy gene that exhibits sequence conservation across different *T. parva* strains (Nene *et al.*, 1992; Nene *et al.*, 1996).

1.1.5 Control of *T. parva*

The current control strategy for controlling *T. parva* is based principally on the elimination of the tick vector by the use of acaricides. Although effective if rigorously applied, the requirement for the continuous use of the acaricides combined with their high cost and often poor availability in rural areas, the emergence of acaricide resistance in tick populations and concerns regarding the contamination of the environment, make continued reliance on intensive application of acaricides for the control of theileriosis unsustainable. Therefore, the search for alternative and more sustainable control measures is a necessity.

1.1.5.1 Immunization by the infection and treatment protocol

The use of the I & T protocol for immunisation using either a 'cocktail' of mixed parasite strains (e.g. the 'Muguga cocktail' used by Radley *et al.* (1975b) which comprises 3 *T. parva* strains - Muguga, Serengeti-transformed and Kiambu 5) or a parasite isolate that is known to contain heterologous strains (e.g. *T parva* (Marikebuni) used by Morazaria *et al.* (1988)) has been shown to be both economical and effective across a spectrum of different farming systems under 'in-field' conditions (Minjauw *et al.*, 1999; Morzaria *et al.*, 1988; Nyangito *et al.*, 1994; Pegram *et al.*, 1996; Young *et al.*, 1992). The successful use of sporozoite strains in areas geographically distant from where they were first isolated indicates that the antigenic composition of *T. parva* strains is not significantly different across its geographical distribution (Mutugi *et al.*, 1989; Oura *et al.*, 2004).

Although applied extensively in some areas, for example Eastern Zambia, where >150,000 calves were immunized by the I & T protocol between 1987 and 2002 (Marcotty *et al.*, 2002), I & T immunization has not attained widespread use due to a variety of drawbacks. Prominent amongst these are the lack of local infrastructure for the production and quality-control of sporozoites and the cold-chain on which distribution of viable sporozoites for immunisation is reliant. Additionally, because it gives rise to a carrier status (Kariuki *et al.*, 1995; Maritim *et al.*, 1989), there is also a

perceived risk of introducing foreign immunisation strains into native tick populations previously free of them (Geysen *et al.*, 1999).

1.1.5.2 Subunit vaccines

The development of a subunit vaccine is generally considered to offer the best long term solution to the control of ECF (McKeever and Morrison, 1998; Morrison, 1996b).

Since 1992 various forms of a neutralizing subunit vaccine based on recombinant p67 sporozoite surface antigen have been developed and assessed under laboratory conditions (Bishop *et al.*, 2003; Heussler *et al.*, 1998; Kaba *et al.*, 2005; Musoke *et al.*, 1992; Nene *et al.*, 1995) with the results indicating that ~70% of vaccinates are protected against severe disease after experimental challenge with an LD_{70} dose of sporozoites. However, only a small proportion of immunised animals achieve complete neutralisation of infection, with the remainder developing schizont parasitosis of varying severity upon challenge. As all vaccinates develop high titres of neutralizing antibody, it is inferred that the variation in protection afforded by p67 vaccination may be attributable to other immune mechanisms which have not been quantified in the trials (Bishop *et al.*, 2003; Musoke *et al.*, 1992). Recent results from the first field trial of a p67 subunit vaccine reported that the efficacy of vaccination averaged only ~50% for a group of 97 animals located at 3 different sites in Kenya that were exposed to field challenge for a period of up to 4 months following vaccination (Musoke *et al.*, 2005).

Given the protective role of $CD8^+$ T-cells in response to *T. parva*, CTL epitopes are obvious candidates for inclusion in subunit vaccines. By immuno-screening APCs transiently transfected with *T. parva* (Muguga) schizont cDNA clones, using CTL derived from I & T immunised animals of diverse MHCI haplotypes, 6 antigens (Tp1, Tp2, Tp4, Tp5, Tp7 and Tp8) containing CTL epitopes have been identified (Graham *et al.*, 2006). Tp1 and 2 have been annotated as hypothetical proteins, whereas Tp4, 5, 7 and 8 were predicted to encode the ε-subunit of T complex protein 1, elongation translation factor 1A, heat-shock protein 90 and cysteine protease,

respectively (Gardner *et al.*, 2005). These antigens were used to vaccinate two groups of 12 cattle employing either plasmid DNA/ modified vaccinia virus Ankara strain (MVA) or canary pox/MVA primer/boost vaccination regimes. (Graham *et al.*, 2006). As there were no significant differences between the two vaccination regimes the authors presented the results as a single group. Although 19 out of the 24 vaccinated cattle (79%) developed antigen-specific $CD8^+$ T-cell IFN-γ responses (as measured by an ELISpot assay), only 7 of these animals produced a CTL response. On subsequent challenge with a lethal dose of homologous sporozoite only these 7, and an additional 3 vaccinates (~40% of the vaccinated animals) survived (Graham *et al.*, 2006). Moreover, all of the vaccinates developed clinical disease when challenged. This disappointing result may reflect the failure of the antigen-delivery strategies to elicit antigen-specific $CD8^+$ T-cells with appropriate function for protection and/or the failure of the vaccination protocols to induce adequate $CD4^+$ T-cell responses in vaccinates (Graham *et al.*, 2006). As discussed above, there is evidence that the latter are required to provide help for the activation of CTL (section 1.4.4).

In the future, generation of a multi-component subunit vaccine against *T. parva* incorporating p67, CTL and $CD4^+$ epitopes able to provoke the relevant effective responses would present the ideal solution. However, as in humans, the major challenge for subunit vaccination in cattle remains the development of appropriate antigen-delivery systems that reproducibly evoke potent CTL responses (Graham *et al.*, 2006; McKeever *et al.*, 1999).

1.1.5.3 Immunodominance and subunit vaccination

A major issue that will need to be addressed in the use of $CD8^+$ T-cell antigens in *T. parva* subunit vaccines is the effect of immunodominance. If subunit vaccination does not achieve subversion of the immunodominance hierarchy then the immunity induced will remain focused on a limited number of polymorphic immunodominant epitopes and as such display limited parasite strain cross-reactivity. In some murine and primate systems subject to immunodominance, subunit vaccination has been shown to be capable of concurrently eliciting CTL to both immunodominant and

subdominant epitopes (Chen *et al.*, 1998; Fu *et al.*, 1997; Santra *et al.*, 2002), and thereby induce potentially increased breadth of protection. However, evidence from other immunodominance models suggests that it may be necessary for the immunodominant and subdominant epitopes in the subunit vaccine to be delivered separately to ensure presentation by different APCs in order to subvert immunodominance (section 1.3.2.4 - Kedl *et al.*, 2000; Pion *et al.*, 1999; Wolpert *et al.*, 1998) and generate more cross-protective responses.

An alternative strategy may be to identify more broadly conserved subdominant epitopes to include in a subunit vaccine (Altfeld and Allen, 2006). It is known that CTLs specific for subdominant epitopes have a role in controlling viral infections, such as HIV, where there is immunodominance amongst the CTL epitopes (Frahm *et al.*, 2006); and subunit vaccination using subdominant epitopes has proven to afford protection against SIV in rhesus monkeys (Santra *et al.*, 2002) and influenza in mice (Oukka *et al.*, 1996).

Identification of *T. parva* antigens recognised by $CD8^+$ T-cells on defined MHC backgrounds (Graham *et al.*, 2006) provides for the first time the opportunity to quantitatively assess immunodominance in the *T.parva*-specific $CD8^+$ T-cell response. As clonal diversity and the TCR repertoire of T-cells responding to antigens has been shown in other systems to be an important factor in both the efficacy of $CD8^+$ T-cell responses and their ability to recognise antigenic variants (section 1.2.6.1) it was also of interest in this project to examine these features of the *T. parva*-specific $CD8^+$ T-cell response. Therefore, in the remainder of the introduction, background information on TCR (in particular the TCRβ chain) and immunodominance is discussed.

1.2 The αβ T-cell receptor

An essential feature of the adaptive immune system of jawed vertebrates is its ability to discriminate between the vast array of antigens it encounters and then mount appropriate antigen-specific responses. This is achieved by the expression of diverse repertoires of clonally distributed antigen-specific receptors (AgRs) on lymphocytes: immunoglobulin (Ig) on B-cells and αβ and γδ T-cell receptors (TCR) on αβ and γδ T-cells respectively. Ig and TCRs share many characteristics, including genomic organisation of TCR and Ig genes and tertiary protein structure (Chothia *et al.*, 1988; Davis and Bjorkman, 1988), and are thought to have evolved from a common ancestor, although the exact Ig/TCR evolutionary history remains unresolved (Glusman *et al.*, 2001; Richards and Nelson, 2000).

The αβTCR is a disulphide-linked heterodimer composed of an α and β polypeptide chain, each of which consist of (i) a short cytoplasmic tail, (ii) a hydrophobic transmembrane domain, (iii) a short hinge region containing a cysteine residue that forms the inter-chain disulphide bond, (iv) a membrane-proximal constant (C) domain and (v) a membrane-distal variable (V) domain. On the cell surface TCRs are associated with invariant accessory chains of the CD3 complex, which function in signal transduction following TCR ligation.

αβTCRs recognise antigen in the form of short processed peptides presented in association with MHC molecules on the surface of APCs. αβT-cells are categorised into two major classes on the basis of effector function and expression of either $CD4^+$ or $CD8^+$ MHC binding co-receptors. The ligands for αβTCRs of $CD8^+$ 'cytotoxic' T-cells are peptide-MHCI complexes (pMHCI) and the ligands for αβTCRs of $CD4^+$ 'helper' T-cells are pMHCII. MHCI molecules generally bind peptides derived from degraded intra-cytosolic proteins whilst MHCII molecules bind peptides derived from exogenous proteins that have been degraded in the lysosomal pathway.

1.2.1 Somatic recombination in antigen-specific receptors

A distinguishing characteristic of AgRs is that exons encoding the variable domains of the component polypeptide α, β, γ, δ TCR chains and IgH (heavy) and Igκ/λ (Funkhouser *et al.*) immunoglobulin chains are not present in the germline but are formed during lymphocyte development by somatic recombination of variable (V), diversity (D - β, δ and IgH only) and joining (J) gene segments (Tonegawa, 1983). The chromosomal loci for each of the chains (TCRB for TCRβ chain, TCRG for TCRγ chain, TCRA/D for TCRα and δ chains, IGH for IgH, IGK for Igκ and IGL for Igλ) contain multiple copies of these genes; the different permutations of V(D)J combination that can be achieved for individual lymphocytes provides an integral source of diversity for AgR repertoires (section 1.2.4).

Recombination signal sequences (RSS), which are composed of a conserved heptamer and nonamer sequence separated by a spacer of either 12bp (12-RSS) or 23bp (23-RSS), are found at the 3'end of V segments, 5'end of J segments and both the 3'end and 5' end of D segments (Hesse *et al.*, 1989; Max *et al.*, 1979; Ramsden *et al.*, 1994; Sakano *et al.*, 1979). These RSS are recognised by lymphocyte-specific recombination activating protein 1 (RAG1) and RAG2, which initiate somatic recombination by creating double stranded breaks that cleave the RSS from gene segments undergoing recombination (McBlane *et al.*, 1995). The gene segments are then recombined through the action of ubiquitously expressed DNA repair machinery (Weaver, 1995). Recombination is restricted by the 12/23 rule; by which recombination can only occur between a gene segment flanked by a 12-RSS and one flanked by a 23-RSS (Tonegawa, 1983). The 12/23 rule is sufficient to guarantee correct recombination of V(D)J elements, except for TCRβ and δ chains, where direct V gene (12-RSS) to J gene (23-RSS) recombination is avoided by mechanisms 'beyond the 12/23 rule' which ensure the interposition of a D gene between the V and J segments (Bassing *et al.*, 2000; Hughes *et al.*, 2003; Jung *et al.*, 2003; Olaru *et al.*, 2005).

V(D)J recombination is tightly controlled in a lineage- and stage-specific way so that although all Ig and TCR chain recombination uses the same generic RAG1/2 and

RSS system, Ig and TCR are only rearranged in B-cells and T-cells respectively and rearrangement of loci occurs only at the correct developmental stages, e.g. rearrangement of the TCRB locus prior to the TCRA locus in αβ T-cells (reviewed in Bassing *et al.*, 2002; Schatz, 2004; Schlissel, 2003). Strict regulation of V(D)J recombination is also essential in enforcing allelic exclusion in which all loci, except TCRA, are limited to the production of a single functional rearrangement (reviewed in Khor and Sleckman, 2002; Mostoslavsky *et al.*, 2004). In contrast, the TCRA locus is allelic inclusive and as a consequence about 30% of αβ T-cells contain mRNA transcripts for two functionally rearranged TCRα chains and thus can express two distinct αβ TCRs that share the same β chain (Borgulya *et al.*, 1992; Casanova *et al.*, 1991; Heath and Miller, 1993; Padovan *et al.*, 1993).

1.2.2 Nomenclature of TCR genes

Two alternative standardised systems of TCR gene nomenclature have been established by the WHO-IUIS Nomenclature sub-committee on TCR designation (Kazatchkine, 1995) and by the IMmunoGeneTics database (IMGT - Giudicelli and Lefranc, 1999). The major conceptual difference between the two systems is in the nomenclature of V gene segments. In both systems V genes are organised into subfamilies (WHO-IUIS)/subgroups (IMGT) whose members share over 75% nucleotide similarity in germline configuration. In the WHO-IUIS nomenclature, subfamilies (and members of subfamilies) are numbered arbitrarily according to their order of discovery, whereas in the IMGT system subgroups (and members of subgroups) are numbered based on their position in the loci starting from the 5'end.

Due to differences in (i) the order of discovery of V genes and (ii) evolution of the loci in mouse and human, orthologous murine and human V genes rarely have corresponding subfamily/subgroup names in either system. Tables that detail the correspondence between the two nomenclature systems for V genes are available through the IMGT website (http://imgt.cines.fr) and the table for the human TCRβV genes has been reproduced in Appendix A. In both the IMGT and WHO-IUIS

systems it is convention that V genes of other species are designated subfamily numbers corresponding to the human V gene with which they share highest nucleotide similarity (Baron *et al.*, 2001; Butler *et al.*, 2005; Houston and Morrison, 1999; Isono *et al.*, 1994; Schrenzel *et al.*, 1994). Following its approval by the Human genome organisation (HUGO) nomenclature committee, the IMGT system is considered official and is widely used, but the WHO-IUIS nomenclature is still cited in many current publications (e.g. Gillespie *et al.*, 2006; Trautmann *et al.*, 2005) and will be used in this study to provide continuity with previously published work on bovine TCR genes (Buitkamp *et al.*, 1993; Houston *et al.*, 1999; Tanaka *et al.*, 1990)

The notation used in the WHO-IUIS system is predominantly self explanatory. Following the TCR prefix, initials are used to described the locus (A,B,G,D) and gene segment type (V,D,J,C). For V genes the succeeding number identifies the subfamily, whereas for the D, J and C genes this number identifies, if appropriate, the DJC cluster in which the genes are located (section 1.2.3). The subsequent S number is used in V genes to enumerate and distinguish subfamily members; for example, TCRBV5s4 refers to T-cell receptor gene of TCRB locus, variable gene subfamily 5, member 4, which can be further abbreviated to Vβ5s4. In J gene nomenclature the S number is used to delineate the genes within a cluster, which are numbered according to genomic order proceeding from the 5'end.

1.2.3 The human and murine TCRB loci and TCRβ genes

1.2.3.1 The TCRB locus

The genomic sequences of all TCR loci in human and mouse are now available (GenBank accession numbers are given in Appendix B) and have provided an important resource for studying various aspects of TCR gene biology. The TCR loci are large, incorporating hundreds of kilobases (Kb), ranging from nearly 200Kb for the human TCRG locus to over 1.7Mb (Megabases) for the murine TCRA/D locus. Within each TCR locus there is an array of V gene segments at the 5'end and a region containing the (D)-J-C elements at the 3' end. The TCRB locus is located on

chromosome 7q34 in humans and chromosome 6A-C in mice (Barker *et al.*, 1984; Caccia *et al.*, 1984; Isobe *et al.*, 1985) and is approximately 650Kb in length in both species (Rowen *et al.*, 1996, GenBank accession numbers AE000663-665).

Schematic representations of the human and murine TCRB loci are given in Figure 1.4. The TCRB locus has three distinctive characteristics:

(i) The duplication of the entire DJC region, leading to the presence of two tandemly located DJC clusters, each consisting of a single 5' Dβ gene, six to seven Jβ genes and a single 3' Cβ gene. The Jβ genes are approximately 50bp in size and are flanked at their 5' end by a 12-RSS. The Dβ genes are very short, 9-16bp, and flanked by 12-RSS and 23-RSS on their 5' and 3' ends respectively. The Cβ genes are present as four exons which have a combined length of ~550bp.

(ii) The presence of a solitary Vβ gene (Vβ20s1 in humans and Vβ14s1 in mice) that is located 3' to the DJC clusters and lies in the reverse transcriptional orientation to the other TCRβ genes. It has been hypothesized that the anomalous location and orientation of this gene is due to its exclusion from an ancient inversion event that incorporated the rest of the TCRB locus (Glusman et al 2001) and

(iii) The interposition of trypsinogen genes 5' and 3' to the Vβ gene array. This association is ancient (>300 million years), as demonstrated by its conservation in the chicken (Glusman *et al.*, 2001).

1.2.3.2 Germline human and murine Vβ repertoire

The human TCRB locus contains 65 Vβ genes distributed amongst 30 subfamilies, and the murine locus 35 Vβ genes distributed amongst 31 subfamilies. Each Vβ gene is composed of two exons; a leader (L) exon, generally ~49bp long, separated by an intron of ~80-250bp from a variable (V) exon of ~300bp which is immediately flanked at its 3'end by a 23-RSS (Siu *et al.*, 1984).

Vβ genes follow a 'birth and death' model of evolution whereby new genes are created by duplication and some are maintained in the genome whilst others are deleted or become non-functional (Clark *et al.*, 1995; Su and Nei, 2001). Phylogenetic trees of human and murine Vβ genes exhibit branching patterns (Arden *et al.*, 1995a, b), indicating that Vβ gene duplications have occurred continuously over a prolonged period of time. Extensive intermingling of human and murine Vβ sequences in phylogenetic trees (Clark *et al.*, 1995; Su and Nei, 2001) reflect the occurrence of many Vβ duplications prior to rodent/primate divergence 80-110 million years ago (Kumar and Hedges, 1998; Li *et al.*, 1990).

Figure 1.4 Schematic representation of the human and murine TCRB loci (overleaf). The human (upper) and murine (Lopes *et al.*) loci are shown in parallel in three segments running from the 5'end (upper panel) to the 3'end (lower panel). Genes within the loci are shown according to the legend below. Vβ genes represented by two coloured boxes have allelic forms of different functional status. Vβ gene names in red type are examples of genes within duplicated cassettes (section 1.2.3.2). Red lines indicate orthologous human and murine Vβ genes as documented in the literature (section 1.2.3.3 - Bosc and Lefranc, 2000; Clark *et al.*, 1995; Glusman *et al.*, 2001; Lai *et al.*, 1988). The arrows adjacent to human Vβ20s1 and murine Vβ14s1 indicate the opposite transcriptional orientation of these genes compared to the other TCRβ genes in the TCRB loci. The boxes representing genes, and the distances between the genes are not shown to scale. Exons are not shown. The format of the figure follows that of the IMGT database (http://imgt.cines.fr).

Legend

(i) Vβ gene segments
- Functional
- ORF pseudogene
- Pseudogene

(ii) Other TCRβ genes
- Dβ gene
- Jβ gene cluster
- Cβ gene

(iii) Non-TCRβ genes
- Trypsinogen genes

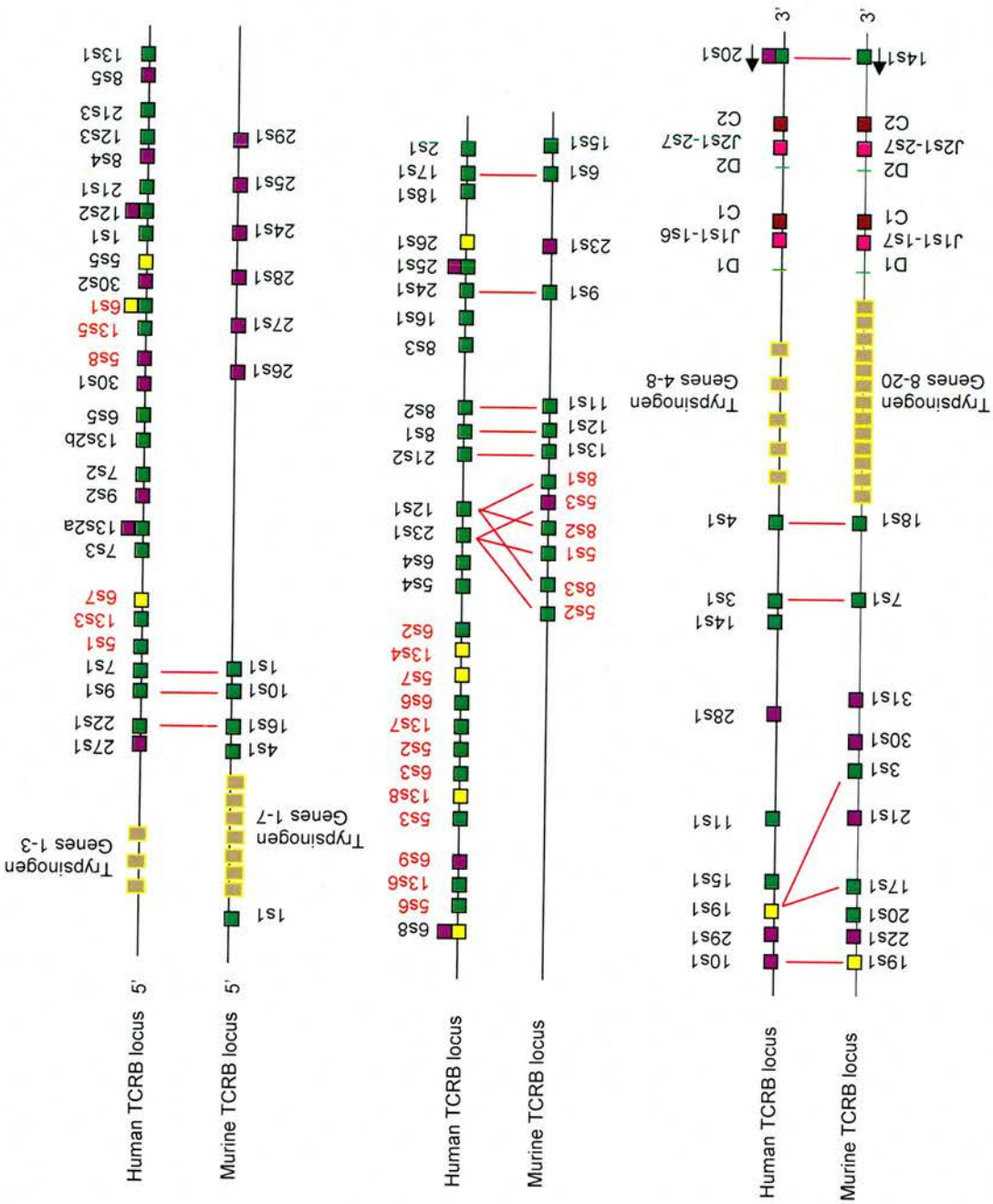


Figure 1.4. Schematic representation of the human and murine TCRB loci. See previous page for description.

The disparity in the number of Vβ genes in the human and murine germline repertoire testifies to more extensive duplication in the human; nine of the subfamilies have multiple members, with the largest subfamilies being Vβ5 (8 members), Vβ13 (9 members) and Vβ6 (9 members). In contrast only two of the murine subfamilies have multiple members - Vβ5 and Vβ8, which have three members each. The exact mechanisms leading to duplication are unknown, but unequal cross-over, possibly facilitated by genomic repeats in the loci (such as retroviral or LINE repeats) has been suggested as a potential mechanism (Glusman *et al.*, 2001). Much of the expansion of the human Vβ gene repertoire has occurred through the duplication of cassettes containing multiple Vβ genes; the most prominent example being the five repeats of a 20Kb region of DNA that includes a Vβ5, 13 and 6 gene (Figure 1.4 - Rowen *et al.*, 1996; Su and Nei, 2001). Similarly the expansion of murine Vβ5 and 8 appears to have occurred following duplication of a cassette containing the ancestral genes of these two subfamilies - leading to three adjacently located pairs of tandemly organised Vβ5/Vβ8 in the middle of the locus (Figure 1.4 - GenBank accession numbers AE000663-665).

1.2.3.3 Vβ pseudogenes and the functional Vβ repertoire

The germline Vβ repertoires of humans and mice include numerous pseudogenes. Nineteen of the human Vβ genes (~30%) in the TCRB locus, and fourteen of the murine Vβ genes (~40%) are non-functional. Some of these pseudogenes lack an open-reading frame (either due to frame-shifts or substitutions creating premature stop codons), whilst others maintain an open-reading frame (ORF pseudogenes) but are non-functional due to defects in splice sites, defects in their RSS that prohibit somatic recombination or nucleotide substitutions that cause the loss of structurally important amino acids (Rowen *et al.*, 1996, GenBank accession numbers AE000663-665). Due to the presence of pseudogenes, not all of the Vβ subfamilies in humans and mice contain functional genes. In humans the functional Vβ repertoire comprises 46 genes in 23 subfamilies and the functional murine Vβ repertoire consists of 21 genes in 19 subfamilies (Bosc and Lefranc, 2000; Folch and Lefranc, 2000b).

In addition to the Vβ genes in the TCRB locus, in humans duplication and inter-chromosomal translocation has generated a cluster of seven Vβ genes present on chromosome 9 (Robinson *et al.*, 1993; Rowen *et al.*, 1996). Because no Dβ, Jβ or Cβ genes were incorporated in the translocation, these genes are unable to form TCRβ chains and are therefore termed 'orphan' pseudogenes.

1.2.3.4 Orthology between human and murine Vβ genes

Various criteria have been used to describe orthology between human and murine Vβ genes, including (i) >65% nucleotide sequence identity (Lai *et al.*, 1988), (ii) >60% amino acid sequence identity (Clark *et al.*, 1995), (iii) proximity on phylogenetic trees (Clark *et al.*, 1995; Su and Nei, 2001) and (iv) sharing of distinctive features, e.g. human Vβ23 and murine Vβ5 have L exons of 79bp, rather than the typical ~50bp (Glusman *et al.*, 2001). However, due to complexities introduced into the human and murine TCRB loci by deletion, duplication, mutation and gene conversion events subsequent to human/murine divergence, identification of orthologues amongst the Vβ genes is not always straightforward. For example in some instances, a single subfamily in one species appears to have several orthologues in the other species (e.g. human Vβ19 could be an orthologue of either murine Vβ20, 17 or 3), presumably as a result of gene duplication in the ancestral locus and subsequent deletion of genes in one of the species, or alternatively duplication of an ancestral gene in the other species subsequent to divergence (Clark *et al.*, 1995; Glusman *et al.*, 2001).

Such complications and the application of different criteria have led to different researchers producing discordant results when describing human/murine Vβ orthologues. Thus, Clark *et al.* (1995) identified 14 subfamily orthologies, whereas Bosc and LeFranc (2000) identified only 11. Orthologous relationships between murine and human TCRβ genes that have been described in the literature are shown in Figure 1.4, and demonstrates the conservation of order of orthologous Vβ genes between the human and murine TCRB loci (Lai *et al.*, 1988).

1.2.4 Diversity in the αβTCR repertoire

1.2.4.1 Mechanisms of generating αβTCR diversity

A diverse αβTCR repertoire is integral to the ability of T-cells to mount efficient antigen-specific responses (reviewed in Goldrath and Bevan, 1999; Nikolich-Zugich *et al.*, 2004). Diversity of the αβTCR repertoire is generated during assemblage of TCRs in thymocytes (developing T-cells) by two mechanisms; combinatorial diversity and junctional diversity (Davis and Bjorkman, 1988).

Combinatorial diversity occurs at two levels: (i) different permutations of V(D)J genes in re-arrangements used to form α and β chains and (ii) different combinations of rearranged α and β chains in αβTCRs. The amount of diversity that can be produced is a function of the number of functional Vα, Jα, Vβ, Dβ and Jβ genes available for recombination - in humans combinatorial diversity theoretically could produce ~3 x 10^5 unique αβTCRs.

Junctional diversity dramatically amplifies this diversity by editing the nucleotide sequence of the V(D)J junction through the activities of exonuclease and terminal deoxynucleotide transferase (Tdt), which respectively remove and insert bases to the 5'end of V, the 5' and 3' ends of D and the 3'end of J gene segments undergoing somatic recombination. The accumulation of combinatorial and junctional diversity can theoretically produce a potential repertoire of 1 x 10^{15} unique αβTCRs (Davis and Bjorkman, 1988), which is several orders of magnitude greater than the ~1-2 x 10^8 and 10^{12} T-cells present in mice and human respectively (Arstila *et al.*, 1999; Casrouge *et al.*, 2000). As a consequence of the mechanisms used, much of the diversity represented in this repertoire is focused within the V(D)J regions (also known as the CDR3 regions - section 1.2.5) of the α and β chains.

1.2.4.2 Diversity of the naïve αβTCR repertoire

Estimates of the actual diversity of the peripheral αβTCR repertoire present in individual mice and humans have been obtained by extrapolation of direct molecular measurements (Arstila *et al.*, 1999; Casrouge *et al.*, 2000). The results suggest that

the naïve human and murine repertoires comprise in the order of 2 x 10^6 and 2.5 x 10^7 unique αβTCRs respectively; thus the αβTCR repertoire expressed by individuals represents a fraction of the potential repertoire that can be generated. Various factors have been demonstrated to influence which αβTCRs will constitute an individuals repertoire, the most significant of which are (i) thymic selection, (ii) non-random rearrangement and (iii) germline polymorphisms of the TCRB locus.

1, Thymic selection

Once αβTCRs are expressed in the thymus, thymocytes undergo positive and negative selection before emigrating to join the mature peripheral pool (reviewed in Starr *et al.*, 2003; von Boehmer *et al.*, 2003). The affinity of αβTCRs for pMHC expressed on thymic epithelial and dendritic cells determines the fate of individual thymocytes during these selection processes. In positive selection, thymocytes bearing αβTCR with low but appreciable affinity for self pMHC are rescued from 'death by neglect', which is the fate of thymocytes bearing αβTCRs with no appreciable affinity for self pMHC. Negative selection subsequently deletes thymocytes bearing αβTCR with high affinity for pMHC from the repertoire. Thus the combined effect of positive and negative selection is to produce mature T-cells with a repertoire of αβTCRs that are capable of interacting with self-MHC molecules but are not auto-reactive. Less than 5% of thymocytes survive selection, dramatically reducing the diversity of αβTCR expressed by T-cells exported from the thymus (Correia-Neves *et al.*, 2001; Nikolich-Zugich *et al.*, 2004; Sant'Angelo *et al.*, 1997).

As MHC proteins are highly polymorphic, for example, in man more than 800 alleles of MHCI and more than 600 alleles of MHCII proteins are known (Robinson *et al.*, 2003), the MHC ligands that participate in thymic selection differ between individuals and result in variation in the αβTCRs that survive thymic selection. Human cohort studies have demonstrated that individuals sharing MHC haplotypes have more similar peripheral αβTCR repertoires than individuals with different MHC haplotypes (Akolkar *et al.*, 1993; Gulwani-Akolkar *et al.*, 1991; Mizushima *et al.*, 1997); and in MHC disparate mice strains, the αβTCR repertoires also show markedly different frequencies of expression of certain Vβ subfamilies (Bill *et al.*,

1988). Similarly, the thymic selection of $CD4^+$ and $CD8^+$ T-cells against MHCI and MHCII proteins respectively has been observed to generate significant differences in the frequency at which these T-cell subsets express certain Vβ genes (Hawes *et al.*, 1993; Liao *et al.*, 1990). Evidence from several studies has shown that the influence of MHC phenotype on αβTCR repertoire is mediated through both positive (Benoist and Mathis, 1989; Blackman *et al.*, 1989; Liao *et al.*, 1989; MacDonald *et al.*, 1988a) and negative selection (Bill *et al.*, 1989; Kappler *et al.*, 1987; Liao *et al.*, 1990).

Another factor that can influence the αβTCR repertoire surviving thymic selection is the presence of endogenous super-antigens (eSAGs). eSAGs are molecules encoded by retroviruses integrated into the genome that can bind to MHCII molecules outside of the peptide binding groove to form a ligand for αβTCR expressing certain Vβ genes (Herman *et al.*, 1991). eSAG interaction with immature thymocytes leads to clonal deletion of all cells bearing αβTCRs expressing the relevant Vβ genes (Kappler *et al.*, 1988; MacDonald *et al.*, 1988b; Meylan *et al.*, 2005; Nelson *et al.*, 2004; Pullen *et al.*, 1988); for example, expression of the eSAG Minor Lymphocyte Stimulating antigen 1a (Mls-1a) in some strains of mice causes clonal deletion of $V\beta8s1^+$ and $V\beta6^+$ T-cells (Kappler *et al.*, 1988; MacDonald *et al.*, 1988b).

2, Non-random rearrangement

Thymocyte populations that have not undergone thymic selection demonstrate biased expression patterns of Vβ and Jβ genes (Ema *et al.*, 1997; Jores and Meo, 1993; Manfras *et al.*, 1999; Nanki *et al.*, 1998) due to preferential usage of certain TCRβ genes in somatic recombination. These biases are maintained in post-selection repertoires (Ema *et al.*, 1997; Nanki *et al.*, 1998; Rosenberg *et al.*, 1992) and therefore have substantial impact on mature peripheral αβTCR repertoires.

Although the RSS associated with different Vβ genes is generally conserved, there is variation in the sequences of the heptamer, nonamer and spacer components of RSS (Glusman *et al.*, 2001; Ramsden *et al.*, 1994). Extensive analysis of Ig recombination has shown that RSS sequence variation can determine the frequency with which V genes are used (Feeney *et al.*, 2004; Montalbano *et al.*, 2003; Nadel *et al.*, 1998a;

Nadel *et al.*, 1998b). Similarly, *ex vivo* and *in vitro* studies have demonstrated that the frequency with which certain TCRβ genes are used in rearrangement is also influenced by the RSS (Livak *et al.*, 2000; Wu *et al.*, 2003) and it has been postulated that use of certain orthologous Vβ genes at strikingly similar frequencies in humans and mice may reflect conservation of their RSS sequences (Livak, 2003, 2004)

As each Vβ gene is under the control of its own promoter and there are substantial sequence differences between individual Vβ gene promoters (Halle *et al.*, 1997), disparity in promoter activity has also been cited as a possible explanation for preferential use of specific Vβ genes (Chen *et al.*, 2001a; Ryu *et al.*, 2004; Sieh and Chen, 2001). Unlike biased use of Ig V genes, biases in the usage of Vβ genes is not dependent on their proximity to the DJC region (Sieh and Chen, 2001; Wilson *et al.*, 2001).

3, Germline polymorphisms in the TCRB locus

Various murine TCRB locus germline deletions have been described in laboratory mouse strains and in wild mice (Behlke *et al.*, 1986; Haqqi *et al.*, 1989; Jouvin-Marche *et al.*, 1989; Kotzin *et al.*, 1985; Noonan *et al.*, 1986; Pullen *et al.*, 1990). A 8.8Kb deletion described in New Zealand White mice incorporates the Cβ1, Dβ2 genes and the Jβ2 cluster of genes (Kotzin *et al.*, 1985; Noonan *et al.*, 1986) whilst the other deletions reported involve loss of various Vβ elements which can involve >50% of the entire Vβ germline repertoire (Haqqi *et al.*, 1989). Two insertion/deletion-related polymorphisms (IDRPs) have been reported in the human TCRB locus (Seboun *et al.*, 1989), one of which involves three Vβ genes (Zhao *et al.*, 1994).

Some of the human and murine Vβ genes have allelic variants (Bosc and Lefranc, 2000; Folch and Lefranc, 2000b). For some, e.g. human Vβ20s1 and murine Vβ17s1, there are non-functional as well as functional alleles (Charmley *et al.*, 1993; Wade *et al.*, 1988). For other Vβ genes, alleles are found to be expressed at different levels in the αβTCR repertoire due to polymorphic differences in the coding region (Vissinga

et al., 1994), the promoter region (Kay *et al.*, 1994), the RSS sequence (Posnett *et al.*, 1994) or interactions with super-antigens (Cazenave *et al.*, 1990).

The loss of functional TCRβ genes either by germline deletion or presence of non-functional allelic forms and the existence of alleles that are expressed at different frequencies will obviously impact on the level of expression of genes within the naive αβTCR repertoire.

4, Stochastic selection

Despite the constraints placed on an individual's αβTCR repertoire by germline polymorphisms, non-random rearrangements and thymic selection, evidence from two murine studies suggests that a large redundancy of diversity is maintained. In one study, analysis of the naive repertoire of two genetically identical mice revealed that only 20-25% of the TCRβ chains expressed were shared (Bousso *et al.*, 1998); whilst in another study, the reconstituted TCR repertoires of several alymphoid mice that received spleen cells from the same donor, were found to be highly divergent, with up to 80% of the TCR sequences being unique to individual recipients (Maryanski *et al.*, 2001). Thus, it would appear that the naïve repertoire of an individual is also influenced by the stochastic expression of αβTCR from an enormous available repertoire.

1.2.5 αβTCR interaction with pMHC ligands

Over the last decade high resolution analyses of 24 αβTCRs bound to pMHC antigens have provided detailed structural information of how αβTCRs interact with their ligands (reviewed in Garcia and Adams, 2005; Hennecke and Wiley, 2001; Housset and Malissen, 2003; Rudolph *et al.*, 2006). As predicted from comparison to immunoglobulins (Chothia *et al.*, 1988; Davis and Bjorkman, 1988) the αβTCR antigen-binding surface is composed of 6 loops, called complementarity determining regions (CDRs), that are brought together at the membrane distal end of the receptor as shown in Figure 1.5 (Garboczi *et al.*, 1996; Garcia *et al.*, 1996). The α and β chains each contribute 3 loops (CDRα 1-3 and CDRβ 1-3 respectively). The CDR1s

and CDR2s are encoded within the germline V genes and the CDR3s are encoded by the V(D)J junctional region. Consequently, as most of the nucleotide diversity of the TCRα and β chains is focused within the V(D)J region, the CDR3 regions display a much greater level of variability than the CDR1 and CDR2 regions (section 1.2.4.1).

In all of the αβTCR/pMHC structures described there is a conserved diagonal positioning of the αβTCR across the pMHC ligand which maximises TCR/pMHC interaction (Garcia and Adams, 2005; Rudolph *et al.*, 2006).This orientation places the germline encoded CDR1 and 2 loops predominantly in contact with the conserved α-helices of the presenting MHC molecules, whilst the highly variable CDR3 loops primarily interact with the peptide, as shown in Figure 1.5. However, there is considerable variation within this generalised orientation of TCR/pMHC structures; the angle at which the TCR lies on the pMHC can vary by up to ±45° (Rudolph *et al.*, 2006; Teng *et al.*, 1998) and the tilt of the TCR on the pMHC also displays variability, such that in some complexes the CDRβ1 and 2 loops make minimal contact with the MHC (Marrack *et al.*, 2001; Teng *et al.*, 1998). Consequently the contribution made by the different CDR loops to the binding energy varies between individual TCR/pMHC complexes (Borg *et al.*, 2005; Manning *et al.*, 1998).

1.2.6 Antigen-specific TCR repertoires

Each pMHC ligand is recognised by a repertoire of antigen-specific αβTCRs selected from within the naïve repertoire. Experimental evidence that TCRβ chain transgenic mice, in which the diversity of the naïve TCR repertoire is markedly decreased, can mount normal responses against a panel of antigens indicates that there is a large redundancy of TCRs available for recognising specific epitopes (Listman *et al.*, 1996). However, identification of 'holes' (i.e. pMHC for which there are no cognate TCR) in the TCR repertoire subsequent to loss of certain TCR genes in other studies shows there are limits to this redundancy (Kumar and Sercarz, 1994; Nanda *et al.*, 1991).

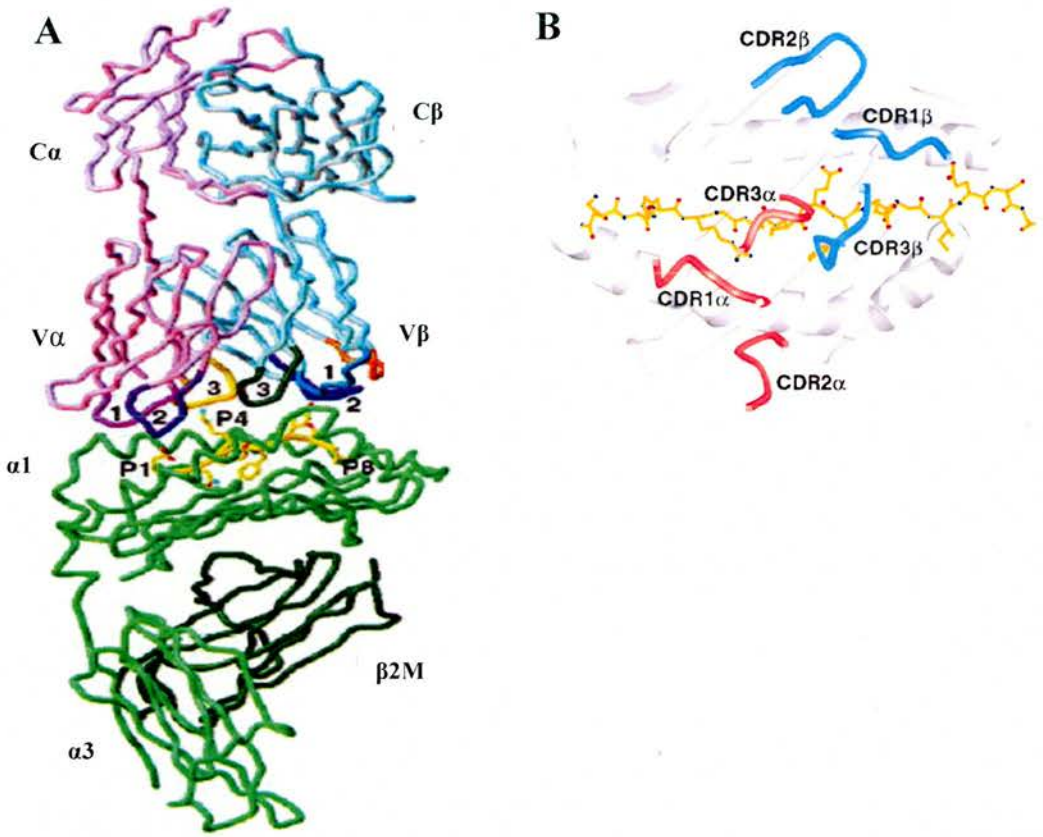


Figure 1.5 TCR interaction with pMHC ligand. (A) Lateral view of a typical TCR/pMHC complex. TCR on top (α chain in pink, β chain in light blue), pMHC on bottom (MHCI heavy chain - α1, α2 and α3 helices in light green and β_2 microglobulin (β_2M) in dark green, peptide in yellow with amino acid positions 1, 4 and 8 (P1, P4 and P8) indicated). TCR binding surface is formed by six complementarity determining regions (CDRs). CDR1α magenta, CDR2α purple, CDR3α yellow, CDR1β cyan, CDR2β navy blue and CDR3β green (Garcia *et al.*, 1998). (B) CDR placement over the composite peptide-MHC ligand in a typical TCR/pMHC complex. MHC helices shown as white ribbons, peptide in yellow with the NH2- (N) and COOH- (C) termini indicated, TCRα chain CDRs in red and TCRβ chain CDRs in blue (adapted from Hennecke and Wiley 2001).

The available naïve αβTCR repertoire obviously has a significant impact on the composition of antigen-specific repertoires (Attuil *et al.*, 2000; Bousso *et al.*, 1998).

Germline deletion of TCR genes (Brawand *et al.*, 1999; Wallace *et al.*, 2000), expression of allelic variants of TCR genes (Gahm *et al.*, 1991) and negative selection against co-expressed MHC products (Belz *et al.*, 2000; Burrows *et al.*, 1995) have all been documented as altering antigen-specific TCR repertoires. However, the stochastic selection of αβTCRs from within the available naïve repertoire also plays a role in moulding the shape of antigen-specific repertoires (Attuil *et al.*, 2000; Bousso *et al.*, 1999; Kedzierska *et al.*, 2004).

1.2.6.1 Diverse and restricted TCR usage in antigen-specific TCR repertoires

Some antigen-specific repertoires are extremely diverse (Campos-Lima *et al.*, 1997; Casanova *et al.*, 1991; Horwitz *et al.*, 1994; Taylor *et al.*, 1990; Utz *et al.*, 1996), but for many pMHC the specific repertoire is restricted, showing use of particular Vα and/or Vβ genes (and hence conserved CDR1 and 2 sequence) in combination with conservation of the hypervariable CDR3 sequence (Argaet *et al.*, 1994; Cose *et al.*, 1995; Gillespie *et al.*, 2006; Lehner *et al.*, 1995; Moss *et al.*, 1991).

Various factors have been identified as influencing the diversity of antigen-specific TCR repertoires. (i) Maryanski and colleagues proposed that epitopes that bear close homology to self-peptides have restricted repertoires due to limitations in the diversity of cognate αβTCR that survive negative selection (Casanova *et al.*, 1992; Casanova and Maryanski, 1993; Maryanski *et al.*, 1997). (ii) Evidence from experimental models suggest that the stability of pMHC complexes is correlated with the ability to elicit clonal expansion of both high and low avidity T-cells and that therefore stable pMHC are able to generate a more diverse antigen-specific TCR repertoire than unstable pMHC (Busch and Pamer, 1998; Campos-Lima *et al.*, 1997). (iii) Results from two recent structural studies indicate that the topology of the composite pMHC ligand may influence the antigen-specific TCR repertoire, with featureless or super-bulged pMHC complexes inducing limited diversity, whilst pMHC with prominent features selected a diverse TCR repertoire (Miles *et al.*, 2005; Turner *et al.*, 2005). (iv) The diversity of an antigen-specific repertoire can also be restricted by the presence of a dominant TCR clonotype that exhibits optimal binding

to the pMHC ligand (Kjer-Nielsen *et al.*, 2002; Stewart-Jones *et al.*, 2003). For example, in the case of the influenza peptide FLRGRAYGL presented by HLA-B8, the dominant TCR clonotype is cross-reactive with the alloantigen HLA-B*4402 and therefore deleted in HLA-B8$^+$/HLA-B*4402$^+$ individuals - in these individuals the FLRGRAYGL/HLA-B8 antigen-specific repertoire is diverse compared to that seen in HLA-B8$^+$/HLA-B*4402$^-$ individuals (Burrows *et al.*, 1995).

Diversity of antigen-specific TCR repertoires in CD8$^+$ T-cells responses has been associated with enhanced anti-viral protection in both human and mouse models (Lopes *et al.*, 2003; Messaoudi *et al.*, 2002). This has been attributed to greater TCR diversity allowing, (i) selection of the most efficient CD8$^+$ effector T-cells (Messaoudi *et al.*, 2002) and (ii) enhanced ability to recognise epitope variants through the presence of multiple 'fine specificities' within a diverse antigen-specific TCR repertoire, and therefore control of 'escape mutants' in viral pathogens such as HIV (Douek *et al.*, 2002; Lopes *et al.*, 2003; Pantaleo *et al.*, 1997).

1.2.6.2 Stability of antigen-specific TCR repertoires

Once established antigen-specific repertoires are not fixed, and although some clones may persist for up to 57 months (Levitsky *et al.*, 1998; Maini *et al.*, 2000) other clones does not display such stability and are lost from the repertoire whilst novel clones become incorporated into it (Annels *et al.*, 2000; Chen *et al.*, 2001c). Repeated exposure to antigen can significantly alter the TCR repertoire through the process of 'immuno-focusing' - the narrowing of TCR diversity (Kedzierska *et al.*, 2004; Maryanski *et al.*, 1999; McHeyzer-Williams and Davis, 1995). Analysis of immuno-focusing in several bacterial and viral models indicate that this narrowing is the result of preferential selection of T cells with high affinity TCR (Busch and Pamer, 1999; Savage *et al.*, 1999; Trautmann *et al.*, 2005; Zhong and Reinherz, 2004). Based on results from *in vivo* studies Kedl *et al.* (2002) have proposed that this selection of high affinity TCR is mediated by the ability of T-cells bearing such TCR to out-compete T-cells with low affinity TCR for pMHC presented by APC. Consequently, T-cells expressing high affinity TCR are preferentially expanded

during antigenic challenge and more likely to contribute to the memory cell pool that responds in subsequent exposures.

1.2.6.3 Characteristics of CD8+ T-cell responses

As will be discussed in section 1.3 many $CD8^+$ T-cell responses to complex antigens are focused on a limited number of immunodominant epitopes. Studies using tetrameric pMHCI complexes have demonstrated that *in vivo* $CD8^+$ T-cell responses to the immunodominant epitopes of viruses can be massive. During the primary response to Epstein-Barr virus (EBV) infection, 44% of $CD8^+$ T-cells in one individual were specific for a single dominant epitope (Callan *et al.*, 1998) and even in individuals that have recovered from EBV infection >5% of the circulating $CD8^+$ T-cell population are specific for the dominant epitope (Callan *et al.*, 1998; Tan *et al.*, 1999). Similarly large expansions of antigen-specific $CD8^+$ T-cells have been demonstrated in the response to LCMV infection in mice (Butz and Bevan, 1998).

Adoption of techniques that permit characterisation of the αβTCR of large numbers of antigen-specific T-cells has shown that the antigen-specific repertoire of $CD8^+$ responses are frequently polyclonal but dominated by a small number of large oligo-clonal expansions (Chen *et al.*, 2001c; Maryanski *et al.*, 1996; Naumov *et al.*, 1998; Peggs *et al.*, 2002). Another feature of some $CD8^+$ antigen-specific TCR repertoires is the occurrence of 'public' TCRs, i.e. identical antigen-specific TCRs present in different individuals sharing the same MHC haplotype (Argaet *et al.*, 1994; Levraud *et al.*, 1996; Quinn *et al.*, 2006; Zhong and Reinherz, 2004). Structural studies have demonstrated that two such 'public' TCRs of humans exhibit optimal binding to their pMHC ligands, mainly attributable to a combination of features unique to these particular clonotypes (Kjer-Nielsen *et al.*, 2002; Kjer-Nielsen *et al.*, 2003; Stewart-Jones *et al.*, 2003). The high affinity that 'public' TCRs have for their ligands is thought to account for their recurring preferential selection from the naïve αβTCR repertoire of unrelated individuals (Kedzierska *et al.*, 2004; Kjer-Nielsen *et al.*, 2002; Stewart-Jones *et al.*, 2003; Trautmann *et al.*, 2005).

1.3 Immunodominance

Despite the presence of enormous numbers of potentially immunogenic peptides in viruses, bacteria, tumours and tissue grafts, the bulk of the responding $CD8^+$ T-cell population is typically focused on a limited number of peptides by a phenomenon known as immunodominance (Yewdell and Bennink, 1999). These epitopes are generally categorised into hierarchies of 'immunodominant' and 'subdominant' epitopes, according to the magnitude of the T-cell response they elicit. Within immunodominant $CD8^+$ T-cell responses, the hierarchy of epitopes is usually reproducible in individuals sharing the same MHCI haplotype.

Immunodominance poses a major hurdle to the development of vaccines that aim to induce $CD8^+$ T-cell responses (Yewdell and Del Val, 2004). The focusing of the response against a limited number of epitopes favours the emergence of 'escape mutants' in viruses such as HIV/SIV (Allen *et al.*, 2000; Barouch *et al.*, 2002; Goulder *et al.*, 1997) and limits the breadth of protection afforded against pathogens where the protective $CD8^+$ T-cell response is focused on a limited number of immunodominant epitopes that vary between pathogen strains (e.g. *T. parva* - sections 1.1.3.2 and 1.1.5.2). Consequently, vaccination strategies must aim to modify or circumvent immunodominance to enable the induction of $CD8^+$ T-cell responses to either a diversity of epitopes or broadly conserved epitopes, so that the immunity generated will, where relevant, cover a wide range of pathogen strains and limit the emergence of viral 'escape mutants' (Altfeld and Allen, 2006; Liu *et al.*, 2006; Walker and Korber, 2001).

1.3.1 The MHCI-CD8+ T-cell antigen-presentation pathway

Immunodominance results from the interplay of a multitude of interdependent factors that exert influence at the various stages of the MHCI-$CD8^+$ T-cell immuno-surveillance system (Yewdell and Bennink, 1999; Yewdell and Del Val, 2004).

Peptides presented by MHCI molecules to $CD8^+$ T-cells are predominantly derived from cytosolic proteins degraded by the proteasome, a multi-catalytic protease complex in the cytosol (reviewed in Rock *et al.*, 2002). The proteasome exists in two forms; the constitutive proteasome, which is expressed in most cells, and a modified form known as the immunoproteasome, which is expressed in dendritic cells and cells exposed to IFNγ (Groettrup *et al.*, 1995; Morel *et al.*, 2000). Substitution of the catalytic subunits Y, Z and MB1 in the constitutive proteasome for their counterparts LMP-2, LMP-7 and MECL-1 in the immunoproteasome results in different cleavage site preferences and therefore generation of different peptide pools by the two forms of proteasome (Eleuteri *et al.*, 1997; Tenzer *et al.*, 2004; Toes *et al.*, 2001).

Peptides generated by the proteasome are subsequently conveyed to the endoplasmic reticulum (ER) by the transporter associated with antigen processing (TAP), a heterodimeric membrane-spanning protein (Monaco *et al.*, 1990; Neefjes *et al.*, 1993; Spies and DeMars, 1991). Although proteasomal cleavage dictates the C terminus of antigenic peptides, the N terminus may be extended by several amino acid and require trimming by amino-peptidases (Craiu *et al.*, 1997; Mo *et al.*, 1999); both cytosolic and ER-localised amino-peptidases that trim MHCI epitopes have recently been identified (Rock *et al.*, 2004; Saric *et al.*, 2002; Serwold *et al.*, 2001; Serwold *et al.*, 2002; Stoltze *et al.*, 2000; York *et al.*, 2002). In the ER, the processed peptides are loaded into nascent MHCI molecules and then exported to the cell surface for interaction with $CD8^+$ T-cells.

1.3.2 Factors influencing immunodominance

The mechanisms underlying immunodominance have been the subject of intensive investigation and numerous contributory factors have been identified. These can be categorised broadly into factors that influence (i) abundance of epitope generated, (ii) efficiency of epitope presentation by MHC, (iii) the available T-cell population bearing cognate TCR and (iv) competition between $CD8^+$ T-cells for APCs. As immunodominance is the product of the summation of these factors, it cannot be

definitively attributed to any single factor. Furthermore, the relative importance of the different factors in determining immunodominance varies between different epitopes and systems (Chen *et al.*, 2000; Deng *et al.*, 1997; Pion *et al.*, 1999; Sandberg *et al.*, 1998).

1.3.2.1 Epitope abundance

A correlation between the quantity of epitope generated from its parent protein and its ability to elicit a $CD8^+$ T-cell response has been demonstrated in many models (Del Val *et al.*, 1991; Eisenlohr *et al.*, 1992; Ossendorp *et al.*, 1996; Shastri *et al.*, 1995). However, epitope abundance is not correlated to immunodominance in other models, presumably due to the more significant counter-influence of other factors (Anton *et al.*, 1997; Crotzer *et al.*, 2000).

The sequence flanking an epitope can dramatically alter its liberation from the parent protein (Del Val *et al.*, 1991; Eggers *et al.*, 1995) and consequently its immunogenicity (Moudgil *et al.*, 1998; Vijh *et al.*, 1998). Several recent studies have demonstrated the contribution that the flanking sequences of epitopes can make to immunodominance hierarchies. By exchanging the flanking sequences of an immunodominant and subdominant epitope from ovalbumin, Mo *et al.* (2000) demonstrated that the relative hierarchy of the these epitopes was correlated to the flanking sequences, whilst in another study Zhu *et al.* (2005) demonstrated that the crypticity (i.e. failure to elicit a detectable $CD8^+$ T-cell response) of some epitopes was directly attributed to their flanking sequences. Altered flanking sequences have also been revealed to cause loss of immunogenic epitopes, thus causing a change in immunodominance hierarchies (Beekman *et al.*, 2000; Milicic *et al.*, 2005).

Similarly, the proteolytic activities of the proteasome and amino-peptidases have been shown to affect epitope abundance (Gileadi *et al.*, 1999; Schwarz *et al.*, 2000; York *et al.*, 2002). The efficient generation of some epitopes requires expression of the immunoproteasome (Schultz *et al.*, 2002; Sewell *et al.*, 1999; Sijts *et al.*, 2000), whilst other epitopes are more efficiently generated by the constitutive proteasome (Morel *et al.*, 2000); thus the differential expression of the two proteasomal forms

has the ability to influence immunodominance hierarchies (Basler *et al.*, 2004; Chen *et al.*, 2001b). A recent study using a mutant mouse lacking one of the amino-peptidases that trims MHCI epitopes (endoplasmic reticulum aminopeptidase 1 - ERAP1) demonstrated that absence of this enzyme drastically alters immunodominance hierarchies, thereby indicating the role it has in mediating immunodominance (York *et al.*, 2006).

1.3.2.2 Epitope presentation by MHCI

The crucial influence that peptide affinity for the presenting MHC molecule has on immunodominance has been well documented (Chen *et al.*, 1994; Sette *et al.*, 1994; van der Most *et al.*, 1996). However, in subsequent studies it was appreciated that the immunogenicity of an epitope was more accurately reflected by its ability to form stable pMHC complexes (Levitsky *et al.*, 1996; van der Burg *et al.*, 1996). In all studies in which pMHC stability has been assessed it has been found to correlate with immundominance (Brooks *et al.*, 1998; Bullock *et al.*, 2000; Busch and Pamer, 1998; Gallimore *et al.*, 1998). This has been related to the formation of stable immunological synapses between APCs and $CD8^+$ T-cells, which may have to be maintained for >10hr to achieve full T-cell activation (Davis *et al.*, 2003; Yoshimura *et al.*, 2004). The extended pMHC/TCR dwell time facilitated by stable pMHC complexes may enable maintenance of the immunological synapse for the required period, whereas unstable pMHC may not (Kalergis *et al.*, 2001). Stable pMHC may also contribute to immunodominance by enabling the activation of both high and low avidity $CD8^+$ T-cells whereas unstable pMHC are more likely to be capable of only activating high avidity $CD8^+$ T-cells (Busch and Pamer, 1998).

1.3.2.3 Available T-cell population bearing cognate TCR

It has been well established that the naïve $CD8^+$ T-cell precursor frequency plays an important role in determining immunodominance in some systems (Cao *et al.*, 1996; Choi *et al.*, 2002; Deng *et al.*, 1997; Sandberg *et al.*, 1998; Thatcher *et al.*, 2000). Consequently, as a dominant influence on shaping the naïve $CD8^+$ T-cell repertoire, thymic selection can exert substantial influence over immunodominance hierarchies

(Slifka *et al.*, 2003) and may account for the effect that co-expressed MHC haplotypes can have in altering them (Belz *et al.*, 2000; Lacey *et al.*, 2003). An example of this is the depression of the $CD8^+$ T-cell response to the immunodominant $H2D^b$-restricted influenza polymerase$_{224-233}$ peptide (D^bPA_{224}) in mice that co-express $H2D^k$. This is associated with the loss of a prominent D^bPA_{224}-specific $V\beta7^+$ T-cell response and is thought to reflect $H2D^k$ mediated clonal deletion of these cells during negative selection (Belz *et al.*, 2000).

The influence of the naïve TCR repertoire in moulding immunodominance has been most strikingly demonstrated in experimental models where the immunodominant antigen-specific repertoire is restricted to T-cells bearing specific Vβ genes. For example, the immunodominant epitope of HLA-CW3 recognised by $CD8^+$ T-cells in DBA/2 mice is $CW3_{170-179}$, and the responding T-cells almost exclusively use a specific allele of Vβ10 (MacDonald *et al.*, 1993). In DBA/2 mice that express an alternative Vβ10 allele, an immune response against HLA-CW3 is still generated but the $CW3_{170-179}$ epitope is no longer immunodominant (Bour *et al.*, 1999). Similar changes in immunodominance hierarchies have been observed in other systems where the naïve TCR repertoire is manipulated by alteration of Vβ gene expression (Connolly, 1994; Daly *et al.*, 1995).

The frequency of memory T-cells has an equivalent influence on immunodominance hierarchies. Expansion of the memory $CD8^+$ T-cell pool for subdominant antigens either by antigen-specific vaccination or previous infection with a pathogen sharing a cross-reactive epitope can cause them to assume immunodominance (Brehm *et al.*, 2002; Rodriguez *et al.*, 2002) and in the experimental influenza A model of immunodominance in $H2D^b$ mice, memory $CD8^+$ T-cell frequency plays a crucial role in the assumption of immunodominance by the D^bPA_{224} epitope in secondary immune responses (La Gruta *et al.*, 2006).

1.3.2.4 Competition between CTLs for APCs

Enhancement of the response against subdominant epitopes in the absence of immunodominant epitopes suggests that the T-cell responses against

immunodominant epitopes are capable of inhibiting that against subdominant epitopes (Chen *et al.*, 2000; Deng *et al.*, 1997). Demonstration in a variety of experimental systems that the immunodominance hierarchy could be subverted by the presentation of immunodominant and subdominant epitopes on either an excess of APCs (Grufman *et al.*, 1999; Palmowski *et al.*, 2002; Sandberg *et al.*, 1998) or by segregation on separate APCs (Kedl *et al.*, 2000; Pion *et al.*, 1999; Wolpert *et al.*, 1998) indicated that this inhibition was due to competition at the level of the APC/T cell interaction.

It is known that this competition is an early event under *in vivo* conditions (Willis *et al.*, 2006) but it is still uncertain what mechanism(s) are operating. Competition for cytokines and/or co-stimulatory factors in the APC microenvironments has been proposed (Kedl *et al.*, 2003; Wolpert *et al.*, 1998; Yewdell and Del Val, 2004). In a series of adoptive transfer experiments it was shown that high affinity $CD8^+$ T-cells have a competitive advantage and therefore are more effective at mediating immunodominance (Kedl *et al.*, 2003; Kedl *et al.*, 2002). As with other factors associated with mediating immunodominance, there is evidence that competition for APC may not exert a significant affect in all experimental immunodominant models (Kedl *et al.*, 2003; Probst *et al.*, 2002).

1.4 Aims of the project

In cattle immune to *T. parva*, there is evidence that the $CD8^+$ T-cell response is prominent in mediating protection. Characteristics of this response suggest that it is focused on a limited number of immunodominant epitopes that exhibit polymorphism between parasite strains. Immunodominance amongst polymorphic MHCI-restricted epitopes potentially poses a major challenge to the development of a subunit vaccine for *T. parva,* which will need to elicit effective $CD8^+$ T-cell responses capable of recognising a wide range of parasite strains. Previous studies of the response to *T. parva* were conducted in the absence of any knowledge of the antigens that are recognised by the responding $CD8^+$ T-cells, and evidence of immunodominance was largely inferred from observations on the strain specificity of the responses. The recent identification of target antigens now provides an opportunity to gain a more detailed understanding of immunodominance in the *T.parva* system. This information will contribute significantly to the rational design of a *T. parva* subunit vaccine.

In other systems, epitope-specific $CD8^+$ T-cells responses characterised by diversity in clonal composition and TCR repertoire have been correlated with more effective pathogen control due to increased efficiency and expression of multiple 'fine specificities' permitting recognition of epitope variants. Therefore, characterisation of these two features of epitope-specific $CD8^+$ T-cell responses to potential *T. parva* subunit vaccine candidate antigens is of particular interest.

The principal aims of this study were to; (i) quantify the $CD8^+$ T-cell response to defined MHCI-restricted epitopes and thereby identify immunodominant epitopes that could be candidates for inclusion in a subunit vaccine and (ii) determine the clonal composition and TCRβ repertoire expression of the $CD8^+$ T-cell responses to these epitopes.

In order to address these aims, it was necessary to first develop molecular tools for analyses of bovine TCRβ gene expression. (Chapter 3). The second objective was to

determine the proportion of the $CD8^+$ T-cell responses of immunised animals that were specific for defined epitopes and so provide quantitative data on their relative immunodominance on defined MHC backgrounds (Chapter 6). The third objective was to determine, by the application of the TCR tools developed, the clonal composition and TCRβ chain repertoire expressed by the responding epitope-specific $CD8^+$ T-cells (Chapter 6).

As a consequence of the TCRβ chain sequencing completed in the course of this work, the known bovine TCRβ gene repertoire was substantially expanded by the identification of new genes (Chapter 4). The development of this work also coincided with the generation of bovine genome sequence data. Further examination of the TCRβ gene repertoire by exploration of the incomplete bovine genome revealed an extensive expansion and permitted a partial annotation of the bovine TCRB locus (Chapter 5).

2 Chapter 2 - Materials and methods

2.1 Experimental animals

Animals used in this study were Friesian-Holsteins bred at the Institute for Animal Health (IAH), Compton from parents with serologically defined MHCI phenotypes (Davies *et al.*, 1994). MHC phenotypes of animals 468, 592, 605, 641 and 1011 had been determined using a panel of MHCI-specific mono-clonal antibodies (Ellis *et al.*, 1999); 592 and 1011 were homozygous for the A10 MHCI haplotype, 468 and 641 were homozygous for the A18 MHCI haplotype and 605 was homozygous for the A14 MHCI haplotype.

2.2 Experimental immunization and challenge

The animals used for experimental immunizations were assumed to be naïve as the parasite does not occur in the UK. Immunization was achieved by the infection and treatment protocol as described by Radley *et al.* (1975a). Briefly, animals were infected by subcutaneous administration on the left side of the neck with a lethal dose of cryopreserved Centre for Tropical Veterinary Medicine (CTVM), *Theileria parva* (Muguga) sporozoite stabilate 71, prepared according to the method described by Cunningham *et al.* (1973). Concurrently 20 mg/kg of long-acting oxytetracycline (Terramycin LA, Pfizer, Sandwich, Kent, UK) was administered by deep intra-muscular injection into the gluteal muscles.

In challenge infections, a lethal dose of CTVM *Theileria parva* (Muguga) sporozoite stabilate 71 was administered, as described above (without concurrent oxytetracycline treatment) to animals previously immunised using this stabilate.

The courses of infections were monitored by rectal temperature and assessment of the size of the local draining (i.e. left pre-scapular) lymph node.

2.3 Cellular techniques

2.3.1 Isolation of peripheral blood mononuclear cells (PBMC)

Blood was obtained by jugular venupuncture and collected into an equal volume of Alsever's solution (Appendix C.1) which served as an anti-coagulant and diluent. PBMC were isolated as described in Goddeeris and Morrison (1988). Aliquots of 30 ml of the blood/Alsever's solution mixture were carefully overlaid onto 20 ml of Ficoll-Paque Plus (Amersham Biosciences, Chalford St. Giles, Beds., UK) in 50 ml polypropylene tubes and centrifuged at 900 *x g* for 30 min at room temperature. PBMC were harvested from the Ficoll-Paque Plus/serum interface and mixed with an equal volume of Alsever's solution before pelleting for 10 min at 450 *x g* at room temperature. The pellet was washed three times in Alsever's solution, counted and re-suspended in a suitable volume of the appropriate medium, according to intended subsequent use.

2.3.1.1 Red blood cell (RBC) lysis

If necessary, RBCs contaminating the PBMC were lysed prior to the second wash by re-suspending the pellet in 3 ml of RBC lysis buffer (Appendix C.2) pre-heated to 37°C, and then incubating at 37°C for 1 min.

2.3.2 Antibody-directed, complement-mediated lysis

Cells were re-suspended in standard culture medium (SCM - Appendix C.3) at a final concentration of 1 x 10^7 cells/ml, with 1 ml each of ILA-12 and CC15 hybridoma culture supernatants (section 2.3.2.1) per 5 x 10^7 cells, and incubated on ice for 30 min. Cells were pelleted by centrifugation at 180 *x g* for 10 min at 4°C and washed twice in SCM.

Cells were subsequently re-suspended in SCM at 5 x 10^6 cells/ml, rabbit serum (section 2.3.2.2) added at a ratio of 1 in 5 and the mixture incubated at 37°C for 40 min to allow complement lysis to occur. The cells were pelleted and washed twice in

SCM, re-suspended in 30 ml SCM, layered over 20 ml Ficoll-paque Plus in a 50 ml polypropylene tube, and then centrifuged at 900 *x g* for 20 min at room temperature. Live, unlysed cells were collected from the Ficoll/SCM interface, mixed with an equal volume of SCM and pelleted at 450 *x g* for 10 min at room temperature. The effectiveness of the lysis was assessed by FACS analysis of the remaining live cell population (section 2.3.7).

2.3.2.1 Hybridoma lines

ILA-12 and CC15 (Table 2.1) hybridomas were cultured in SCM at 37°C in a humidified atmosphere of 5% CO_2 until exhausted and the supernatant then harvested. Titration experiments were completed to establish appropriate concentrations for their use (data not shown). The results indicated that a dilution of 1:5 - 1:10 was suitable for both supernatants.

2.3.2.2 Rabbit serum

Rabbit blood was collected by intra-cardiac puncture immediately post-mortem into plain 50 ml polypropylene tubes. The blood was allowed to clot at room temperature for 3 h and then stored at 4°C overnight. Following fragmentation of the clot, the blood was spun at 450 *x g* for 10 min and the serum harvested. Titration experiments were completed to establish the appropriate concentration for serum use as a source of complement (data not shown). The results indicated that a dilution of 1:5 was suitable.

2.3.3 *In-vitro* generation of *T. parva*-specific $CD8^+$ T-cell enriched cell-lines

The protocol for generation of *Theileria parva*-specific $CD8^+$ T-cell enriched cell-lines by *in-vitro* stimulation was adapted from that described by Goddeeris and Morrison (1988).

First stimulation - PBMC were isolated and counted as described above (section 2.3.1). Autologous *T. parva* (Muguga) infected cells (stimulators - section 2.3.5) were harvested, counted and exposed to 50Gy of gamma irradiation from a ^{137}Caesium source for 20 min. Into each well of 24-well plates were placed 2 ml of SCM containing 2 x 10^6 PBMC/ml and 1 x 10^5/ml irradiated autologous stimulators. Plates were incubated for 7 days in a humidified atmosphere of 5% CO_2 at 37°C.

Second stimulation - Cells from the first stimulation were harvested, counted and re-seeded into 24-well plates in 2 ml of SCM at a density of 2 x 10^6/ml together with irradiated autologous stimulators at 1 x 10^5/ml. Plates were incubated for 7 days in a humidified atmosphere of 5% CO_2 at 37°C.

Third stimulation – Cells from the second stimulation were harvested, counted and the $CD4^+$ and $WC1^+$ γδ T-cell populations lysed as described above (section 2.3.2). The remaining cells were re-seeded into 24-well plates in 2 ml of SCM at a density of 1 x 10^5/ml together with irradiated autologous stimulators at a density of 2 x 10^5/ml, irradiated autologous PBMC (fillers - irradiation as for stimulators) at a density of 2 x 10^6/ml and 100U/ml recombinant human IL-2 (Chiron Corporation, Emeryville, CA.,USA), and incubated for 7 days in a humidified atmosphere of 5% CO_2 at 37°C.

2.3.4 Cloning of *T. parva*-specific $CD8^+$ T-cells

Cell lines from third stimulations were harvested and assessed for $CD8^+$ purity by FACS (section 2.3.7). If there was significant contamination (>10%) of the $CD8^+$ population with either $WC1^+$ γδ T-cells and/or $CD4^+$ T-cells, the cells were subject to another round of antibody-directed, complement-mediated lysis (section 2.3.2), until >90% of the population were $CD8^+$. Cells from the third stimulation were then counted and serially diluted in SCM to give a concentration of 40, 20, 10, 5, and 2.5 cells/ml. Aliquots of 100µl of these suspensions were distributed into the wells of 96-well plates to give plates containing 4-0.25 cells/well. To each well was added a 100

µl aliquot of SCM containing 5 x 10^4/ml irradiated autologous stimulators, 2 x 10^5/ml irradiated autologous fillers and 200U/ml of recombinant human IL-2. Plates were incubated for 2 weeks in a humidified atmosphere of 5% CO_2 at 37°C. After 2 weeks plates were visually screened and wells displaying obvious growth were selected for further expansion. Only wells from plates showing growth in <30% of wells were selected, as they have greater than 83% probability of being truly clonal (Henry *et al.*, 1980).

2.3.4.1 Expansion of clones

Selected clones were transferred to individual wells in 48-well plates. To each well was added 1ml of SCM containing 5 x 10^5/ml irradiated autologous stimulators, 5 x 10^5/ml irradiated autologous fillers and 100U/ml of recombinant human IL-2. Plates were incubated in a humidified atmosphere of 5% CO_2 at 37°C, and clones re-stimulated every two weeks as required.

2.3.5 Generation and maintenance of *T. parva*-infected cell lines

Theileria parva (Muguga) transformed (TpM) cell lines were established *in vitro* from PBMC of animals 468, 641, 1011 and 592 as described in Goddeeris and Morrison (1988). In brief, 2 x10^7 PBMC were isolated (as described in section 2.3.1), pelleted by centrifugation at 200 *x g* for 5 min, re-suspended in 0.5 ml of CTVM *Theileria parva* (Muguga) sporozoite stabilate 80 suspension and incubated in a humidified atmosphere of 5% CO_2 at 37°C for 1.5 h with gentle mixing every 15 minutes. The cells were then washed once in 10 ml SCM, re-suspended in 10 ml of SCM, divided into two culture flasks, incubated in a humidified atmosphere of 5% CO_2 at 37°C and visually monitored for transformation of lymphoid cells.

Once established TpM cell lines were kept incubated in a humidified atmosphere of 5% CO_2 at 37°C and maintained in log growth by removal of three-quarters of the cultured cell suspensions and replacement of this volume with SCM every 2/3 days. The phenotype of the TpM cell lines established were determined by fluorescence-

activated cell sorting (FACS) analysis using a panel of mAb specific for CD4, CD8, γδ T-cells, CD3 and IgM as described below (section 2.3.7). 468TpM, 641TpM and 1011TpM had phenotypes consistent with $CD4^+$ αβ T-cells whilst the phenotype of 592TpM was that of a B-cell.

2.3.6 Cryopreservation of cellular material

Cells to be cryopreserved were pelleted, re-suspended at 2 x 10^6 cells/ml in 90% heat-inactivated Foetal bovine serum (FBS)/10% dimethylsulphoxide (DMSO; Sigma, Poole, Dorset, UK) and 1ml aliquots transferred into polypropylene cryovials (CryoTube, Nunc, Roskilde, Denmark). These vials were placed into an isopropanol jacketed container (Cryo 1°C Freezing Container, Nalgene, Neerijse, Belgium) and slowly frozen at -70°C before being transferred to liquid nitrogen storage. Upon use, cryopreserved material was rapidly thawed by incubation in a water bath at 37 °C.

2.3.7 Fluorescence-activated cell sorting (FACS) analysis

The monoclonal antibodies (mAbs) used for primary labelling in FACS analysis are listed in Table 2.1. mAbs derived from culture supernatants were used at a dilution of 1:5 in FACS media (Appendix C.4), while those derived from ascitic fluid were used diluted 1:1,000 in FACS medium.

Aliquots of 50 µl of 2 x 10^7 cells/ml suspensions were distributed in wells of 96-well round-bottom well plates and 50 µl of the required primary labelling mAbs added. Plates were incubated at 4°C for 30 min, washed three times in FACS medium and then re-suspended in 50 µl of fluorescein-isothiocyanate (FITC) labelled goat polyvalent anti-mouse immunoglobulin G,A,M antibody (Sigma-Aldrich, Poole, Dorest, UK). Following incubation at 4°C for 30 min, the cells were washed three times in FACS medium and then re-suspended in 200 µl FACS medium for analysis using a FACScan fluorescence activated cell sorter (Beckton Dickinson, Mountain

View, CA., USA.). Negative controls were incubated with FACS medium instead of primary labelling antibody.

Anitbody	Isotype	Specificty	Cell distribution
IL-A12[a]	IgG2a	CD4	$CD4^+$ T-cells
IL-A105	IgG2a	CD8	$CD8^+$ T-cells, NK-T cells
IL-A51[b]	IgG1	CD8α chain	$CD8^+$ T-cells, NK-T cells
CC15[c]	IgG2a	WC1	Peripheral γδ T-Cells
GB21A[d]	IgG2b	γδ TCR	All γδ T-cells
ILA-24[e]	IgG1	MyD-1	Monocytes/Macrophages/Dendritic cells
ILA-30[f]	IgG1	IgM	B-cells
MM1A[g]	IgG1	CD3	T-cells
NKp46-AKS1[i]	IgG1	NKp46	NK-like cells

Table 2.1 Monoclonal antibodies used for primary labelling in FACS analysis. References defining the specificity of mAbs: Bensaid and Hadam (1991)[a], MacHugh and Sopp (1991)[b], Morrison and Davis (1991)[c], Davis *et al* (1996)[d], Ellis *et al* (1988)[e], Naessens *et al* (1988), Davis *et al* (1993)[g], Davis and Splitter (1991)[h], Storset et al (2004)[i].

2.3.8 Cytotoxicity Assays

^{51}Chromium (^{51}Cr) and ^{111}Indium (^{111}In) 4h-release assays were performed essentially as described by Goddeeris and Morrison (1988).

2.3.8.1 Peptide loading of target cells

Theileria annulata-infected cell lines (TA) 592TA and 641TA were used as 'null' target cells presenting MHCI A10-restricted and MHCI A18-restricted epitopes from *T. parva* (Muguga) respectively. Peptide for the A18-restricted epitope Tp1.1 (VGYPKVKEEML) was kindly provided by Dr. E. Taracha (ILRI, Nairobi, Kenya) and peptides for the A10-restricted epitopes Tp2.1 (QSLVCVLMK) and Tp2.2 (KSSHGMGKVGK) were produced by Pepscan Systems (Lelystad, Netherlands). Immediately before labelling with radioactive isotope, target cells were re-suspended

at 2 x 10^6 in cytotoxicity medium (Appendix C.5) and incubated for 15-30 min with 1 μg/ml peptide.

2.3.8.2 Radioactive labelling of target cells

^{51}Cr and ^{111}In (GE Healthcare Ltd., Little Chalfont, Bucks, UK) were used as the labelling isotopes depending on their availability and the number of effector cells available. Labelling of *T. parva*-infected cells with ^{51}Cr and ^{111}In normally results in the incorporation of 0.1-0.4 and 2-6 cpm/cell respectively (Goddeeris and Morrison, 1988). Thus, approximately 10 times fewer ^{111}In-labelled target cells than ^{51}Cr-labelled target cells are needed for cytotoxicity assays and there is an equivalent reduction in the number of effector cells required.

Target cells were re-suspended at 2 x 10^7 cells/ml in cytotoxicity medium. For labelling with ^{51}Cr, 0.1 ml of the cell suspension (i.e. 2 x 10^6 cells) was incubated with 100 μCi of ^{51}Cr for 1 h at 38°C in a 15 ml polypropylene tube; the cells were then washed three times in 10 ml of cytotoxicity medium before being re-suspended in 2 ml of cytotoxicity medium (i.e. 1 x 10^6 cells/ml). For labelling with ^{111}In, 50 μl of the cell suspension (i.e. 1 x 10^6 cells) was incubated with 5μCi of ^{111}In for 15 min at 38°C in a 15 ml polypropylene tube; the cells were then washed six times in 10 ml of cytotoxicity medium before being re-suspended in 10 ml of cytotoxicity medium (i.e. 1 x 10^5 cells/ml).

2.3.8.3 ^{51}Cr and ^{111}In 4-h release cytotoxicity assays

$CD8^+$ T-cell enriched cell-lines were re-suspended in cytotoxicity medium at a concentration of 1 x 10^7 for ^{51}Cr or 1 x 10^6 for ^{111}In 4 h-release assays. In duplicate, 200 μl/well of two-fold dilutions of the cell-lines were distributed into either 96-well, flat bottom culture plates (for ^{51}Cr release assays) or 96-well V-bottomed culture plates (for ^{111}In release assays) resulting in a range of effector cell concentration of 200 x 10^4 to 6.25 x 10^4 cell/well for ^{51}Cr release assays and 200 x 10^3 to 6.25 x 10^3 cell/well for ^{111}In release assays. Aliquots of 100 μl of un-quantified $CD8^+$ T-cell clone suspensions were used in ^{111}In release assays for clones derived from the cell-

lines; these assays were completed unduplicated in 96-well V-bottomed culture plates.

To each well containing effector cells ($CD8^+$ T-cell-enriched cell-lines or clones) 50 µl of target cell suspension (i.e. 5 x 10^4 ^{51}Cr or 5 x 10^3 ^{111}In labelled cells) were added. In the $CD8^+$ T-cell-enriched cell-line assays this resulted in effector to target cell ratios of 40:1 to 1.25:1.

In triplicate, 50 µl of target cell suspension was added to wells containing (i) cytotoxicity medium - to allow measurement of the spontaneous release of isotope label and (ii) 0.2% (v/v in water) Tween 20 (Sigma, Poole, Dorset, UK) - to measure maximal release of the isotope label. In cytotoxicity assays for the $CD8^+$ T-cell-enriched cell-lines 200 µl of cytotoxicity medium/0.2% Tween 20 was used whereas in the $CD8^+$ T-cell clone assays 100 µl was used.

Plates were centrifuge for 1 min at 180 *x g* and then incubated for 4 h at 38°C in a humidified atmosphere of 5% $C0_2$. Following incubation the plates were again centrifuged for 1 min at 180 *x g* and then half of the volume of the supernatant (carefully avoiding aspiration of any of the pelleted cells) from each well (125µl for assays with $CD8^+$ T-cell enriched cell-lines and 75µl for assays with $CD8^+$ T-cell clones) was transferred to a sample vial and the gamma emissions measured over a period of 1 min in a Wallac Wizard 1470 Automatic Gamma Counter (PerkinElmer, Beaconsfield, Bucks., UK).For each sample the percentage cytotoxicity was calculated as:

$$\text{Cytotoxicity (\%)} = \frac{\text{sample release - spontaneous release}}{\text{maximal release - spontaneous release}}$$

For measurements taken in duplicate/triplicate the arithmetic mean average cytotoxicity was calculated. Based on experience within the laboratory, an arbitrary parameter of $\geq$10% cytotoxicity was used to define significant cytotoxic activity.

2.4 Molecular Techniques

2.4.1 Isolation of RNA

Total RNA was extracted from cell pellets using Tri-Reagent (Sigma-Aldrich, Poole, Dorset, UK) according to the manufacturer's instructions. Cells were re-suspended in approximately 100 µl of PBS (Appendix C.6) and then lysed in Tri-reagent (1 ml per 5 -10 x 10^6 cells) by repeated pipetting. The homogenate was incubated for 5 min at room temperature to allow complete dissociation of nuclear-protein complexes. Chloroform was then added to the homogenate (0.2 ml chloroform per 1 ml of Tri-reagent) and the mixture was vortexed vigorously for 15 s, allowed to stand at room temperature for 2-15 min and then centrifuged at 12,000 *x g* for 15 min at 4°C. The aqueous phase was transferred to a fresh eppendorf, mixed with isopropanol (0.5 ml per 1 ml Tri-reagent), incubated for 5-10 min at room temperature and then centrifuged at 12,000 *x g* for 10 min at 4°C to precipitate the RNA. The RNA pellet was washed twice in 1 ml of 75% ethanol, air dried and then re-suspended in 25 µl of nuclease free water. RNA was stored at -80 °C until use.

2.4.2 Quantification of nucleic acids

Absorbance of light at 260 and 280 nm wavelengths by samples of RNA/DNA solutions diluted in nuclease free water were measured using a GeneQuant spectrophotometer (Biochrom, Cambridge, UK). The concentration and quality of RNA and DNA solutions was estimated based on the assumptions that (i) RNA and DNA have an absorbance reading of 1.0 at 260 nm at concentrations of 40 µg/ml and 50 µg/ml respectively and (ii) optimal quality RNA/DNA solutions have a 260nm:280nm reading ratio >1.8 (Sambrook *et al.*, 1989).

2.4.3 Reverse transcription - complementary DNA (cDNA) synthesis

First strand cDNA was synthesized from RNA using the Reverse Transcription System (Promega, Madison, WI, USA). The manufacturers protocol for priming with

Oligo(dT)$_{15}$ primer was followed with the exception of allowing 1 h instead of 15 min for 42°C incubation. Briefly, 1µg (or multiples thereof) of RNA was incubated at 70°C for 10 min, micro-centrifuged briefly and place on ice until added to the reaction mixture which consisted of 4 µl $MgCl_2$ (25mM), 2 µl 10x Reverse Transcription Buffer, 2 µl dNTP mixture (10mM), 0.5 µl Recombinant RNAsin Ribonuclease Inhibitor, 15u avian myeloblastoma virus reverse transcriptase (AMV-RT), 0.5 µg Oligo(dT)$_{15}$ primer and nuclease free water added to give a final volume of 20 µl. The reaction mixture was incubated at 42°C for 1 h, then incubated at 95°C for 5 min and at 0 °C for 5 min to inactivate the AMV-RT and stop it binding to the cDNA. cDNA was stored at -20°C until use.

2.4.4 Polymerase chain reaction (PCR)

2.4.4.1 Design and synthesis of primers

All primers used (excluding BCON-4, previously designed by Dr. G. Russell and Dr. N. MacHugh) were manually designed during the course of this work. All primers were synthesized by MWG Biotech (Ebersberg, Germany).

(a) Bovine TCRβV(Vβ) subfamily-specific 5' primers

The sequences of the final panel of external (VBext) and internal (VBint) Vβ subfamily-specific primers designed are shown in Table 2.2. Sequences of Vβ subfamily-specific primers that were designed but subsequently rejected are not shown, but are available on request. The number succeeding the (*) denotes the version of the primers adopted for the final panel - from herein, except for where specified, Vβ subfamily-specific primers are described without reference to the version being used and refer to the primers presented in Table 2.2.

(b) Bovine 'Pan-Vβ' primer

A highly degenerate 'Pan-Vβ' 5'primer was designed as described in section 3.3.3. The sequence of this primer was: GGG G(AGCT)(AGCT) (AGCT)(AGC)T (AGT)T(ACT) (CT)TG GTA.

External primer VBext	VBext sequence	Internal primer VBint	VBint sequence
VB1*1ext	TCCCCTGA(AGC)TCTGGA CAC	VB1*1int	GGTACCAACAGGCCCTGG
VB2*1ext	CTCTCGTCTCTCAGCAGC	VB2*1int	GATGCCACATACGAACAAGG
VB3*3ext	CGTGATAGAGTGCATACAGAAC A	VB3*1int	GAT GCT GGC ACG GTT GAG
VB4*2ext	GCTCTCGTCTCTCAAAAGCCAA	VB4*1int	CTACTGCCAATCAGGGCTCCA
VB6*2ext	GG(GC)AGGGGCCAGACTGTGAA T	VB6*1int	CCTATCTGAAGATCCAGCCAG
VB7*2ext	CTCAGATACCAAAATACCTAGTC	VB7*1int	CCAGAGCTCATGTTCATCCAC
VB8*2ext	CACCCAGACCCCTAGGCATGA	VB8*1int	AGCTCCTATTGATGAGTCGG
VB9*2ext	TAGAACATCCTAGAATGTGAGC A	VB9*1int	TGCCAAGTCGTTTCTCACCG
VB10*2ext	AC(CT)CAGAGTCCA(AG)GTCATC TG	VB10*1int	CAAACAGCGATTTTCAGCTGA G
VB13*1ext	CACTCAGGACCCCAGATT	VB13*1int	ACCTGGGACACGGGCTGAG
VB14*2ext	AGGCCCCCTGGA(AT)GCCGAT	VB14*1int	CCAAGACACCGCATCACAGA G
VB15*2ext	CACAGGAAAGAGCACTGTAC	VB15*2int	GTACAATGTCTCTCGTAAGGA G
VB16*5ext	GACC(AC)TATTTCTGGACATGAA TCT	VB16*2int	CCAAAGGCCGATTCACAGCA G
VB17A*1ext	CCCA(CT)GGATGCTGAAATCTA	VB17A*3int	CTGGGCTACAGTGCCATG
VB17B*1ext	CCTGCTCAAAAAGGAAGG	VB17B*1int	CTCTCGTGAAAAGAAGCCTTT C
VB20*1ext	GAACTCAGACCATCCATCAGT	VB20*1int	TCCAGAACTTCAAAGCTTCCA GG
VB24*3ext	CCCAAGATACCTGGTTACCG	VB24*1int	GAAGCCAAGCCAAGCACC
VB28*2ext	GCGCAGGATCTTGGAGACTGA	VB28*1int	CGCCAAGATCCGGGATTTG

Table 2.2 The final panel of bovine Vβ subfamily-specific primers.

(c) Bovine TCRβC (Cβ) specific 3' primers

In all, 5 different Cβ-specific 3' primers were used in this study: BCext and BCint were used in combination with Vβ subfamily-specific primers; PVBrev1 and PVBrev2 were used in combination with the 'Pan-Vβ' primer; and BCON-4 was used in sequencing reactions. The primers were designed using publicly available bovine TCRβC sequence data (GenBank accession numbers D90139 and D90140 - Appendix B), their sequences are detailed in Table 2.3, and their annealing sites to the TCRβC gene are shown in Appendix D.

Primer	Sequence
BCint	GGT CAG CTC CAC GTG GTC
BCext	CGT GAC CTG CTT GTT CAC
BCON-4	CTC TGC TTC CGA GGG TTC
PVBrev1	CAC GTG GTC (AG)GG GTA GAA
PVBrev2	GTG ACC TGC TTC CTG TTC

Table 2.3 The bovine Cβ-specific primers.

2.4.4.2 General PCR protocols

All PCR reactions described below and in other sections of this thesis were completed in PTC-100 thermal cyclers (MJR Inc., Watertown, Mass, USA).

(a) Vβ subfamily-specific semi-nested PCR

Eighteen PCR reactions, one for each Vβ subfamily, were completed for each cDNA template. Each first round reaction used 10 pmol each of BCint (Table 2.3) and the relevant VBext primer (Table 2.2), 0.5 unit BIOTAQ (5units/μl Bioline, London, UK), 2 μl SM-0005 buffer (ABgene, Epsom, Surrey, UK – Appendix C.7), 1μl cDNA diluted 1:5-1:20 in DDW, and nuclease free water to give a final volume of 20 μl. The first round programme was as follows: 94°C for 3 min, 5 cycles of (94°C for 40 s, 60°C for 1 min, 72°C for 30 s), 5 cycles of (94°C for 20 s, 60°C for 1 min, 72°C for 30 s) and a final extension period of 72°C for 3 min.

Each second round reaction used the same components as the first round reaction except that 1µl of the first round product rather than cDNA was used as template and the VBext primers were replaced with the corresponding VBint primers (Table 2.2). The second round programme was as follows: 94°C for 3 min, 25 cycles of (94°C for 15 s, 60°C for 1 min, 72°C for 30 s) and a final extension period of 72°C for 3 min.

(b) Pan-Vβ semi-nested PCR

First round reactions were composed of 200 pmol of the 'Pan-Vβ' primer (section 2.4.4.1), 10 pmol of PVBrev2 (Table 2.3), 0.5 unit BIOTAQ (5units/µl Bioline, London, UK), 2 µl SM-0005 buffer (ABgene, Epsom, Surrey, UK - Appendix C.7), 1µl 1:10 diluted cDNA in DDW, and nuclease free water to give a final volume of 20 µl. Second round reactions had the same constituents but used 1µl of the first round product instead of 1:10 diluted cDNA as template and used PVBrev1 (Table 2.3) instead of PVBrev2 as the 3'primer. The programme for both rounds of amplification was as follows: 94°C for 3 min, 5 cycles of (94°C for 1 min, 57°C for 1 min, 72°C for 1 min), 25 cycles of (94°C for 30 s, 60°C for 1 min, 72°C for 1 min) and a final extension period of 72°C for 10 min.

2.4.5 Agarose gel electrophoresis

PCR products/RE digestion products were analysed by agarose gel electrophoresis using either a midi-gel system (Flowgen, Ashby de la Zouch, Leics., UK) or a large gel system (Owl Separation Systems, Portsmouth, NH, USA). 10 µl aliquots of PCR/RE digestion product were mixed with 2µl of 1 in 6 loading buffer (Appendix C.8) and loaded into a 1.5% agarose gel (Bioline, London, UK), prepared with 1 x TAE (Appendix C.9). Electophoresis was carried out at 120 V for 40 min in the midi-gel system or 200 V for 1 h with the large gel system, using 1 x TAE as running buffer. 1µg of 1Kb Plus DNA Ladder (Invitrogen Ltd., Paisley, UK) was used to enable estimation of product sizes.

Following electrophoresis gels were stained in a bath of Ethidium Bromide (0.5 µg/ml) in 1 x TAE for 30 min before visualisation and storage of gel image with the Gel Doc 2000 system (Biorad, Hercules, CA, USA).

2.4.6 Purification of PCR products

The Wizard PCR Preps DNA Purification System (Promega, Madison, WI, USA) was used for the direct purification of PCR products, according to the manufacturer's instructions. Briefly, 30-300 µl of PCR reaction was added to 100 µl of Direct Purification Buffer and vortexed briefly before addition of 1 ml DNA resin and three further brief vortexes over a 1 min period. The DNA/resin mixture was passed though a mini-column and the DNA/resin attached to the mini-column membrane washed with 80% isopropanol. Following two centrifugations at 10,000 *x g* for 2 min (to dry the resin), 40 µl of nuclease free water was applied to the mini-column, the mini-column incubated at room temperature for 1 min before elution of the purified DNA by centrifugation at 10,000 *x g* for 20 s. DNA not used immediately was store at -20°C.

2.4.7 Generation of plasmid sub-clones

Purified PCR products were cloned into pGEM-T Easy vector system (Promega, Madison, WI, USA) according to the manufacturer's guidelines. Ligation reactions contained 5 µl 2x Rapid ligation Buffer, 1 µl pGEM-T Easy Vector, 25 ng PCR product, 1 µl T4 DNA Ligase (3 weiss units/µl) and nuclease free water to give a final volume of 10 µl and were incubated overnight at 4°C. The following day ligated plasmids/PCR product were transformed into High Efficiency Competent JM109 cells (Promega, Madison, WI, USA). Briefly, 50 µl of JM109 (>10^8cfu/µg DNA) were thawed on ice, added to 2 µl of ligation reaction and then incubated on ice for 20 min before heat-shocking for 45-50 s at exactly 42°C and immediately returning to ice for 2 min. 950 µl of SOC medium (Appendix C.10) at room temperature was

added and the cells then incubated for 1.5 h in a rotary shaking incubator at 37°C and 150 rpm. 100 µl of transformation reaction was plated onto duplicate Luria-Bertani (LB)/ampicillin/IPTG/XGal plates (Appendix C.11) and incubated overnight at 37°C. The following day transformed colonies were identified by blue/white screening.

2.4.7.1 Verifying presence of PCR product inserts in plasmid sub-clones

(a) Method A - PCR verification of presence of insert

Selected colonies were picked using a fine pipette tip, streaked onto gridded LB agar plates containing 100µg/ml ampicillin (Appendix C.11) and incubated for 6 h/overnight at 37°C. Quantities of expanded colonies were harvested, added to 10µl of nuclease free water and denatured by heating at 95 °C for 5 min. Following centrifugation at 13,000 *x g* for 5 min (to sediment the cellular debris), the aqueous phase (containing plasmid DNA) could be subjected to appropriate PCR amplification and the product analysed by electrophoresis as described above (section 2.4.4-2.4.5).

(b) Method B - Restriction Endonuclease (RE) verification of presence of insert

Selected plasmid sub-clones were expanded and DNA prepared as described in sections 2.4.7.2 and 2.4.7.3 below. RE reaction mixtures containing 5 µl of plasmid DNA, 20 units EcoR1 (1 µl of 20,000 units/ml - New England Biolabs, Hitchin, Herts., UK), 1 µl 10x Eco R1 buffer (England Biolabs, Hitchin, Herts., UK), and 3 µl nuclease free water were incubated at 37°C for 1 h and the products of the RE digestion subsequently analysed by electrophoresis as described in section 2.4.5. Presence of a RE product of the anticipated size indicated successful insertion of PCR product into the pGEM-T Easy vector which has EcoR1 cutting sites flanking the insertion site.

2.4.7.2 Expansion and preparation of frozen stocks of plasmid sub-clones

Selected sub-clones were picked using a fine pipette tip and inoculated into 6 ml of LB media containing ampicillin at 100 µg/ml (Appendix C.12) in a sterile 30 ml flask and incubated overnight in a rotary shaking incubator at 37°C and 220 rpm. Aliquots of 750 µl of cultures were mixed with 750 µl of a 50:50 mix of LB media and glycerol (Sigma, Poole, Dorset, UK) and stored at -70°C.

2.4.7.3 Plasmid DNA Preparation – ('Mini-preps')

(a) Method A – Alkali Lysis Method adapted from Sambrook *et al* (1989)

Cultures of plasmid transformed bacteria were prepared as detailed above (section 2.4.7.2). Cells from 1.5 ml of culture were pelleted by centrifugation at 10,000 *x g* for 5 min, re-suspended in 150 µl of Re-suspension Buffer (Appendix C.13a) and then lysed in 150 µl of Lysis Buffer (Appendix C.13b). Following addition of 150 µl of Neutralizing Buffer (Appendix C.13c) cellular debris was pelleted by centrifuging at 15,000 *x g* for 30 min. The supernatant was carefully collected (to avoid disturbance of the cellular debris), transferred to a fresh eppendorf and centrifuged at 15,000 *x g* for 15 min at 4°C to pellet out any contaminating cellular debris. Supernatant was again transferred to a fresh eppendorf, mixed with 1 ml of 100% ethanol and incubated at -20°C for 20 min. DNA was pelleted by centrifuging at 13,000 *x g* for 30 min at room temperature, washed once in 450 µl of 70% ethanol, air-dried, re-suspended in 30 µl of nuclease free water and stored at -20°C until used.

(b) Method B - Wizard Mini-preps DNA Purification System

The Wizard Mini-preps DNA Purification System (Promega, Madison, WI, USA) is a modified form of the alkali-lysis method, using mini-columns rather than ethanol precipitation to isolate DNA, and was used according to manufacturer's instructions. Cells from 1.5 ml of plasmid transformed bacterial culture were pelleted (as described in Method A), re-suspended in 200 µl of Cell Re-suspension Solution and then lysed in 200 µl of Cell Lysis Solution. Following addition of 200 µl of Neutralisation Solution, cellular debris was pelleted by centrifuging at 13,000 *x g* for

5 min, and the supernatant mixed with 1 ml of DNA resin before being passed through a mini-column. After washing the column with 2 ml of Column Wash Solution, the mini-column was centrifuged at 10,000 *x g* for 2 min to dry the DNA/resin on the mini-column membrane. 30-50 µl of nuclease free water was added to the mini-column, and following 1 min incubation at room temperature, the mini-column was centrifuged for 20 s at 10,000 *x g*. Eluted DNA was collected and stored at -20°C until used.

2.4.8 TCRβ chain CDR3 heteroduplex analysis (CDR3β-HDA)

2.4.8.1 Principle of the technique

The bovine CDR3β-HDA method developed in this study is based on that described for humans in Bernardin *et al.* (2003) and is represented schematically in Figure 2.1. Aliquots of cDNA from a sample T-cell population are PCR amplified using 5'primers specific for each Vβ subfamily in concert with the Cβ-specific 3'primer BCext (thus amplifying across the CDR3β region). Mono-clonal DNA 'probes' for each subfamily are generated by PCR amplification of a representative plasmid sub-clone using the same Vβ subfamily-specific primers in concert with BCext which is 5'end labelled with the fluorophore 6-carboxy-fluorescein (FAM). The PCR products from the sample T-cell population are mixed with an excess of the labelled probe from the same Vβ-subfamily, the mixture denatured by heating to 95°C (to separate DNA strands) and then cooled to permit re-annealing. For each subfamily, complementarity between the Cβ (5'extremity) and Vβ (3' extremity) regions facilitates the formation of DNA heteroduplices between strands of the 'probe' and the sample T-cell population amplicons.

The mixtures are then subjected to non-denaturing polyacrylamide gel electrophoresis (ND-PAGE). The differences between the CDR3β sequences of the 'probes' and the TCRβ chains present in the sample T-cell population result in 'bulges' in the heteroduplices that retard their electrophoretic migration during ND-

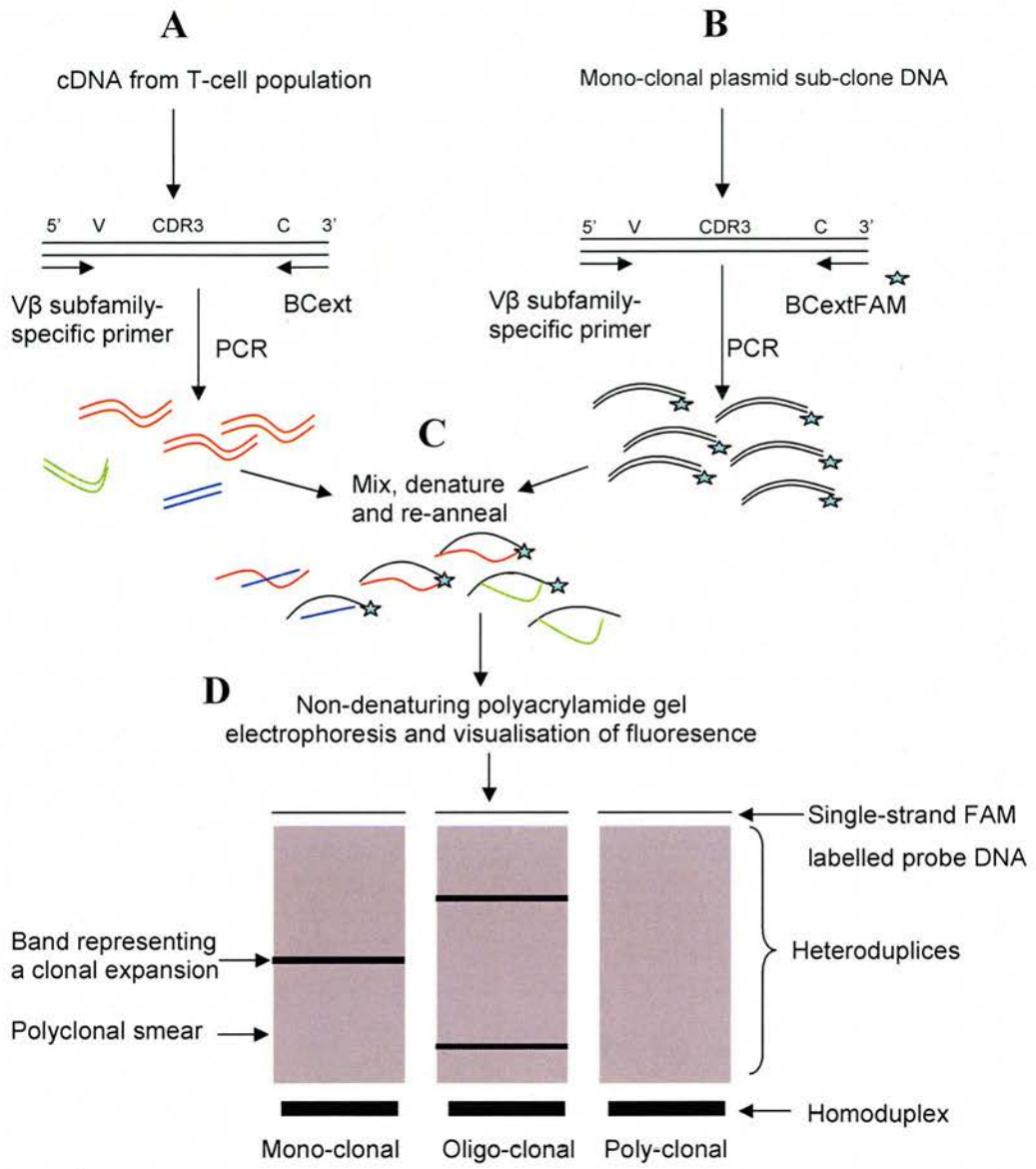


Figure 2.1 Schematic representation of the CDR3β-HDA technique. (A) Aliquots of cDNA from a sample T-cell population are PCR amplified by 5'primers specific for each Vβ-subfamily in concert with the Cβ-specific 3'primer BC-ext. (B) Mono-clonal DNA 'probes' for each subfamily are generated by PCR amplification of a representative plasmid sub-clone using the same Vβ subfamily-specific primers in concert with BCext which is 5'end labelled with the fluorophore FAM (blue star).(C) The PCR products from the sample T-cell population are mixed with an excess of the labelled 'probe' from the same Vβ subfamily, denatured and then re-annealed. (D) The re-annealed mixtures undergo ND-PAGE and are visualized by fluoresence emission. Expanded T-cell clones are represented by distinct bands against the background of a polyclonal smear. Homoduplices of the labelled probe are identified as large bands at the bottom of the gel, whilst single strands of labelled probe DNA may be seen as bands at the top of the gel.

PAGE. As the heteroduplices formed between the 'probe' and the different CDR3β variants in the sample T-cell population will have different structures, the degree of this retardation will vary for the different heteroduplices. Consequently, when the gel is visualised clonal expansions present in the sampled T-cell population appear as distinct bands against the background of a polyclonal smear. Fluorophore labelling of only the reverse strand of the 'probes' results in each clonal expansion being represent as a single band on visualisation by fluorescence emission. Homoduplices of the 'probe' DNA are identified as large bands that run furthest during electrophoresis, whilst single strands of FAM labelled 'probe' DNA may be seen as bands that have migrated very slowly through the gel.

2.4.8.2 Protocol

In eighteen separate reactions, aliquots of cDNA from sample T-cell populations were amplified with Vβ subfamily-specific 5'primers and the communal 3'primer BCext. For subfamilies Vβ1, 10, 14, 16 and 17a the VBint primers (Table 2.2) were used and for all other subfamilies the VBext primers (Table 2.2) were used. Each PCR reaction was composed of 50 pmol each of 5' and 3' primers, 5 µl cDNA, 10 µl SM-0005 buffer (ABgene, Epsom, Surrey, UK - Appendix C.7), 0.5 unit BIOTAQ (5units/µl Bioline, London, UK) and nuclease-free water to give a final volume of 100 µl. The PCR programme was as follows: 94°C for 3 min, 5 cycles of (94°C for 1 min, 60°C for 1 min, 72°C for 1 min), 25 cycles of (94°C for 30 s, 60°C for 1 min, 72°C for 1 min) and a final extension period of 72°C for 10 min.

To generate the FAM-labelled 'probes' for each subfamily aliquots of 25ng of DNA from mono-clonal plasmid sub-clones 1.4, 2.2, 3.1, 4.1, 6.2, 7.2, 8.1, 9.1, 10.1, 13.1, 14.6, 15.1, 16.1, 17A.7, 17B.1, 20.4, 24.1, 28.3 (section 3.3.2) were PCR amplified using the same primers and reaction conditions as for the sample cDNA but with BCext primer 5'labelled with the fluorophore FAM.

A 10 µl aliquot of each sample PCR product was mixed with a 5 µl aliquot of the appropriate Vβ-matched labelled 'probe', the mixture incubated at 95°C for 5 min (to denature DNA duplices) and then at 50°C for 15 min (to allow re-annealing). The

mixtures were kept on ice until loading with 5µl loading buffer (Appendix C.14) onto a 9% non-denaturing polyacrylamide gel (Appendix C.15) in a Biorad Protean® II xi system (Biorad, Hercules, CA, USA). Gels were run at 8 mA for 16 h at 4°C with 0.5 x TBE (Appendix C.16) as running buffer. Gels were visualised and images stored using a Molecular Imager FX (Biorad, Hercules, CA, USA) with a 488 nm laser and a 530 nm band pass emission filter.

For CDR3β-HDA of T-cell clones for which the subfamily of the Vβ gene had already been determined only the relevant reactions were completed and the cDNA used diluted at 1:5 as template for the PCR.

2.4.9 Sequencing

Some PCR products and plasmid inserts were sequenced at either the Moredun Proteomics Facility (Moredun Research Institute, Edinburgh, UK) or DBS Genomics (University of Durham, Durham, UK) following sample preparation by Wizard PCR Preps DNA Purification System (section 2.4.6) or plasmid DNA preparation (2.4.7.3) as appropriate.

However, the majority of sequencing reactions were completed within the laboratory on a 96-well plate format using a BigDye ® Terminator v3.1 (Applied Biosystems, Foster City, CA, USA) protocol adapted from that provided by Lynne Richardson at the Sequencing Unit of the Department of Zoology, University of Oxford.

2.4.9.1 Sample preparation

To each well containing PCR reaction product 60 µl of 20% (w/v) Polyethylene glycol (PEG) MW8000/2.5M NaCl (Appendix C.17) was added, the plate was gently vortexed to ensuring mixing and then incubated at room temperature for 30 min. Plates were centrifuged at 2250 *x g* for 1 h at 4°C to pellet DNA, and the supernatants discarded by centrifuging the inverted plate at 500 *x g* for 1 min. As PEG preferentially binds large DNA fragments centrifugation tends to pellet PCR

products rather than residual primers or dNTPs, which are consequently removed with the supernatants. Following two washes in 150 µl ice cold 70% ethanol, the PCR products were air-dried and re-suspended in a volume of nuclease free water equal to that of the initial PCR reaction. Presence of PCR products and reduction in residual primers was checked by agarose gel electrophoresis as described in section 2.4.5.

2.4.9.2 Sequencing reaction and analysis

Sequencing reactions were composed of 2 µl of purified PCR product, 2.67 pmol of the sequencing primer BCON-4 (Table 2.3), 2 µl 5 x CSA buffer (Appendix C.18), 0.25 µl BigDye ® Terminator v3.1 and nuclease-free water added to a final volume of 10µl. The sequencing reaction programme was 30 cycles of (96°C for 10 s, 50°C for 5 s, 60°C for 2 min).

Sequencing products were prepared for analysis by ethanol precipitation. To each well was added 15 µl nuclease free water, 50 µl 100% ethanol and 2 µl of 3M NaOAc pH5.2 (Appendix C.19). Following mixing by gentle vortex and subsequent incubation at room temperature for 45 min, sequencing products were pelleted by centrifuging for 1 h at 2,250 *x g* at 4°C. The supernatants were discarded by centrifuging the inverted plate for 1 min at 500 *x g*. The DNA pellets were washed once in 150 µl ice cold 70% ethanol and air-dried. Samples were sent to the Sequencing Unit of the Department of Zoology, University of Oxford and analysed using a 3730xl DNA Analyser and Sequencing Analysis software (Applied Biosystems, Foster City, CA, USA).

Where sequencing results contained any ambiguities that could not be resolved by analysis of the sequence chromatogram the sample was re-sequenced and if necessary a consensus sequence formed.

2.5 Sequence Data and Analysis

2.5.1 Nomenclature of bovine TCRβ genes

To provide continuity with previously published work on bovine TCRβ genes (Buitkamp *et al.*, 1993; Houston and Morrison, 1999; Tanaka *et al.*, 1990) the WHO-IUIS nomenclature system (Kazatchkine, 1995) has been used in this study. In accordance with convention, bovine Vβ genes have been given subfamily names according to the human Vβ gene with which they share highest nucleotide similarity (Baron *et al.*, 2001; Houston and Morrison, 1999; Isono *et al.*, 1994; Schrenzel *et al.*, 1994).

Based on observations made on human and murine Vβ genes (Arden *et al.*, 1995a, b) it has been convention to assume (in the absence of contrary genomic data) that Vβ sequences sharing >98% nucleotide sequence identity represent different allelic variants of the same gene, whereas sequences sharing <98% identity are the products of different gene loci. For the analysis of cDNA data and comparison of cDNA and genomic Vβ sequence data this convention has been adopted in this study. Annotation of the part of the bovine TCRB locus incorporating the Dβ, Jβ and Cβ genes (Chapter 5) has permitted the bovine Jβ genes to be named according to their position within the TCRB locus, in accordance with the WHO-IUIS system, and has been applied where appropriate.

To facilitate use of the IMGT/V-QUEST software programme (http://imgt.cines.fr - Giudicelli *et al.*, 2004) the IMGT unique numbering system was applied for description of Vβ genes (Lefranc *et al.*, 2003). This system defines CDR1 and CDR2 as being formed by amino acid residues 27-38 (bp 79-114) and 56-65 (bp 166-195) of the mature Vβ polypeptide chain respectively. Conserved cysteine (Cys/C) residues are located at position 23 and 104, and the tryptophan (Trp/W) at position 41 and the tyrosine (Tyr/Y) at position 42 are also conserved.

2.5.2 Bovine TCRβ gene sequence data

Bovine TCRβ gene sequences deposited in the GenBank database were accessed through the NCBI website (http://www.ncbi.nlm.nih.gov).

The following TCRβV sequences have been submitted by Tanaka *et al.* (1990); D90130 (Vβ1s1), D90121 (Vβ2s1), D90123 (Vβ2s3), D90122 (Vβ2s6), D90127 (Vβ3s1), D90126 (Vβ3s1), D90124 (Vβ4s1), D90133 (Vβ6s1), D90131 (Vβ7s1), D90128 (Vβ15s1), D90129 (Vβ15s2), D90125 (Vβ17s1) and D90132 (Vβ28s1): Buitkamp *et al.* (1993); L18951 (Vβ6s2): and Houston (1997); AJ006346 (Vβ24s), AJ006347 (Vβ20s1), AJ006567 (Vβ2s6), AJ006568 (Vβ1s1), AJ006569 (Vβ1s2), AJ006570 (Vβ1s3), AJ006571 (Vβ1s3), AJ006572 (Vβ1s4), AJ006573 (Vβ1s5), AJ006574 (Vβ1s6), AJ006575 (Vβ1s7), AJ006576 (Vβ1s9), AJ006577 (Vβ14s1), AJ006578 (Vβ13s1), AJ006579 (Vβ10s1), AJ006580 (Vβ6s3), AJ006581 (Vβ2s1), AJ006582 (Vβ2s3), AJ006583 (Vβ2s4), AJ235264 (Vβ10s2), AJ235265 (Vβ10s3), AJ235266 (Vβ13s2), AJ235267 (Vβ13s3), AJ235268 (Vβ16s1).

Additional TCRβV gene sequences were obtained from Houston (1997) for Vβ1s8, Vβ1s10, Vβ1s11, Vβ2s2, Vβ2s5, Vβ2s7, Vβ4s2, Vβ6s3, Vβ7s2, Vβ9s1 and Vβ17s2; and from data made available by Russell and MacHugh (unpublished) for Vβ4s3, Vβ4s4, Vβ8s1, Vβ8s2 and Vβ17s2. The nucleotide and amino acid sequences of these Vβ genes are provided in Appendix E.

Two bovine TCRβC sequences have be deposited by Tanaka *et al.* (1990) - D90139 and D90140. Bovine TCRβJ sequences were obtained from Tanaka *et al.* (1990) and Houston (1997).

2.5.3 TCRβ gene sequence analysis

Details of TCRβ chain cDNA transcript sequence and genomic TCRβ gene sequence analysis are provided in sections 4.2.2 and 5.2.3 respectively. DNAsis v1.0 software (Miriabio Inc., Alameda, CA) was used for completion of most sequence analysis.

2.5.4 Presentation of TCRβ chain sequence data.

The sequences of TCRβ chains have been presented in the standardised format of Chothia *et al.* (1988). The expressed Vβ and Jβ gene segments are identified and the sequences of the 3 framework codons (FR) flanking the CDR3β, as well as the CDR3β itself, are displayed.

3 Chapter Three - Molecular tools for analysing clonal composition and TCRβ repertoire of bovine T-cell responses

3.1 Introduction

The specificity of αβT-cells for particular pMHC ligands is determined by the clonotypically expressed αβTCR (Dembic *et al.*, 1986; Sha *et al.*, 1988). Following recognition of cognate pMHC ligands by engagement of specific αβTCRs, T-cells become activated and undergo clonal expansion. Molecular analysis of TCR expression by T-cells responding to antigenic stimulation has provided insights into many important aspects of T-cell biology including: the clonal composition of antigen-specific T-cell populations (Annels *et al.*, 2000; Bousso *et al.*, 1999; Maini *et al.*, 2000; Maryanski *et al.*, 1999); the diversity of TCR repertoires expressed by antigen-specific T-cells (Gillespie *et al.*, 2006; Lehner *et al.*, 1995; Naumov *et al.*, 1998; Quinn *et al.*, 2006); the development and evolution of T-cell responses (Blattman *et al.*, 2000; Busch and Pamer, 1999; Kedzierska *et al.*, 2004; Sourdive *et al.*, 1998; Zhong and Reinherz, 2004); T-cell responses induced by vaccination (Bennett *et al.*, 2006; Coulie *et al.*, 2001); the relevance of TCR repertoire diversity of responding T cell clonotypes in constraining 'escape mutant' viruses (Douek *et al.*, 2002; Lopes *et al.*, 2003; Pantaleo *et al.*, 1997; Price *et al.*, 2004; Turnbull *et al.*, 2006) and the *in vivo* persistence of T-cell clones in the memory pool (Chen *et al.*, 2001c; Levitsky *et al.*, 1998; Maini *et al.*, 2000; Turner *et al.*, 2003).

In human, primate and murine studies a number of molecular tools have been developed and used to analyze the clonality and TCR repertoires of antigen-specific T-cell responses (reviewed in Rufer, 2005). These techniques have predominantly been developed to analyse the allelically excluded TCRβ chain, although in many studies TCRα chain expression has also been examined. Due to the enormous diversity that is focused into the hypervariable CDR3β region, many techniques are

based on the assumption that each T-cell clonotype expresses a unique CDR3β, and use this as a marker of T-cell clonality.

Immunofluorescence staining of T-cells with Vβ subfamily-specific monoclonal antibodies (e.g. Busch and Pamer, 1999; Cose *et al.*, 1995; Pantaleo *et al.*, 1994; Trautmann *et al.*, 2005) and semi-quantitative RT-PCR with Vβ subfamily-specific primers (e.g. Lehner *et al.*, 1995; Pantaleo *et al.*, 1994) have been widely used to detect alterations in the representation of Vβ subfamilies in responding T-cell populations. However, the information provided by these techniques is of low resolution as although they can be used to identify preferential Vβ subfamily expression, they do not delineate the number of clonal expansions present within expanded Vβ subfamilies or provide information on either the hypervariable CDR3β sequence or Jβ gene usage of expanded clones.

Higher resolution descriptions of T-cell clonality and TCR repertoires have been achieved using various PCR based techniques. TCRβ chain sequencing of antigen-specific $CD8^+$ T-cell clones derived by limiting dilution from antigen-stimulated T-cell cultures has been used in several studies (Cose *et al.*, 1995; Lehner *et al.*, 1995; Utz *et al.*, 1996) and can be combined with functional analysis (Coulie *et al.*, 2001) but is subject to the bias that may be introduced by *in vitro* culture (Dietrich *et al.*, 1997). Modified anchored RT-PCR followed by sub-cloning and sequencing has been used to study the clonality of $CD8^+$ T-cell responses to HIV/SIV infections in humans and primates (Chen *et al.*, 2001c; Douek *et al.*, 2002; Price *et al.*, 2004) and the persistence of adoptively transferred tumour-specific $CD8^+$ T-cells (Robbins *et al.*, 2004; Zhou *et al.*, 2004). This technique provides an unbiased representation of clonality and TCR repertoire but requires a massive sequencing effort. Single-cell RT-PCR of MHCI-tetramer specific $CD8^+$ T-cells has been used in direct *ex vivo* analysis of the clonality/TCR repertoire of responses to defined individual epitopes (Kedzierska *et al.*, 2004; Maryanski *et al.*, 1999; Turner *et al.*, 2003) and provides an elegant and powerful tool for this type of work, although a considerable sequencing effort is also required.

Two extensively used methods that provide a rapid means of detecting expanded T-cell clones in complex mixtures of T-cells are spectratyping (aka immunoscope - Bousso *et al.*, 1998; Cibotti *et al.*, 1994; Levraud *et al.*, 1996; Naumov *et al.*, 1998; Peggs *et al.*, 2002) and TCRβ chain CDR3 heteroduplex analysis (CDR3β-HDA - Bernardin *et al.*, 2004; Lopes *et al.*, 2003; Maini *et al.*, 2000; Turnbull *et al.*, 2006). In spectratyping, distortions of the normal Gaussian distribution of CDR3β length within Vβ subfamilies are taken to reflect clonal expansions (Naumov *et al.*, 1998; Pannetier *et al.*, 1993; Peggs *et al.*, 2002). The CDR3β-HDA technique exploits the retarded migration through non-denaturing poly-acrylamide gel of DNA heteroduplices formed between TCR chains containing Vβ genes of the same subfamily but distinct CDR3β sequences, to identify clonal expansions (section 2.4.8.1). Modifications of this technique have introduced labelled probes that have improved both its interpretation and sensitivity (Bernardin *et al.*, 2003; Maini *et al.*, 1998). In contrast to the other methods, spectratyping and CDR3β-HDA are not dependent on high levels of TCRβ chain sequencing to detect clonal expansions.

At present no molecular tools are available to study the clonal composition or TCR repertoire of bovine T-cell responses. This chapter describes the development of RT-PCR protocols that enable the amplification and sequencing of TCRβ chains of individual bovine T-cell clones and also a CDR3β-HDA technique that is capable of identifying clonal expansions within complex bovine αβT-cells populations.

3.2 Materials and methods

3.2.1 Sequence data and analysis

All available deposited and unpublished bovine Vβ gene sequences (section 2.5.2) were admitted to an 'in-house' database stored using DNAsis v1.0 software (Miriabio Inc., Alameda, CA). Sequence data for bovine Cβ was derived from sequences deposited in the GenBank database (D90129 and D90140 - section 2.4.4.1 and Appendices B and D). Details of TCRβ chain sequence analysis methods are provided in section 4.2.2. TCRβ chain transcript sequences have been displayed in the method defined by Chothia *et al* (1988) as described in section 2.5.4.

The sequence of the artefactual band was compared to bovine and human sequences deposited in the GenBank database using the BLASTN algorithm facility available on the NCBI website (http://www.ncbi.nlm.nih.gov/BLAST).

3.2.2 PCR

3.2.2.1 Vβ-subfamily-specific PCR amplification of cDNA for generation of plasmid sub-clones

Naïve PBMC from animal 605 were isolated (section 2.3.1) and cDNA generated as described in sections 2.4.1-2.4.3. In separate reactions cDNA was amplified using the VBext primers detailed in Table 2.2, with the exceptions noted in Table 3.1, in combination with the common Cβ specific reverse primer BCext (Table 2.3).

Each PCR reaction used 30 pmol of each primer, 1.5 unit BIOTAQ (5units/μl Bioline, London, UK), 6 μl SM-0005 buffer (ABgene, Epsom, Surrey, UK – Appendix C.7), 3μl cDNA, and nuclease free water to give a final volume of 60 μl. The programme used was as follows: 5 min at 94°C, 5 cycles of (1 min at 94°C, 1 min at 60°C, 1 min at 72°C), 25 cycles of (30s at 94°C, 1 min at 60°C, 1 min at 72°C) and a final extension period of 5 min at 72°C.

Vβ subfamily	VBext primer described in Table 2.2	VBext primer used for generation of plamid sub-clone inserts	
		Name	Sequence
3	VBext3*3	VBext3*2	GAT AGA GTG CAT ACA GAA CAT GG
6	VBext6*2	VBext6*1	TTT GCA GGC AGG GGC CAG ACT
8	VBext8*2	VBext8*1	ACA CTT CTC TGC TCA TGC TC
9	VBext9*2	VBext9*1	TAG AAC ATC CTA GAA TGT GAG
14	VBext14*1	VBext14*1	TAT CTC CTG GGA GCA GGC
16	VBext16*5	VBext16*3	GAA TCT GAG TTG TGA CCA TAT
24	VBext24*3	VBext24*1	AGG CTT CTC CTC TGT GTG GC

Table 3.1 Vβ subfamily-specific 5' primers used in the generation of inserts for plasmid sub-clones instead of those described in Table 2.2.

3.2.2.2 Assessing specificity of Vβ-subfamily specific primers

DNA from plasmid sub-clones VB1.1, 2.2, 3.1, 4.1, 6.2, 7.2, 8.1, 9.1, 10.1, 13.1, 14.6, 15.1, 16.1, 17A.6, 17B.1, 20.1, 24.1, 28.3 was prepared by the alkali lysis method (section 2.4.7.3), quantified (section 2.4.2) and diluted in nuclease free double distilled water to give solutions with final concentrations of ~0.5ng/µl.

In individual PCR reactions each of the Vβ subfamily-specific 5' primers, in combination with the common Cβ specific 3'end primer BCint, were used to amplify each of these templates. The final set of Vβ subfamily-specific VBext and VBint 5'primers are detailed in Table 2.2. Each reaction contained 10 pmol of each primer, 0.5 unit BIOTAQ (5units/µl Bioline, London, UK), 2µl SM-0005 buffer (ABgene, Epsom, Surrey, UK - Appendix C.7), 1 µl DNA, and nuclease free water to give a final volume of 20µl. For data relating to the results presented in Figure 3.3 PCRs were completed as described above but using BCext instead of BCint as the 3'primer. The PCR programme used was as described in 3.2.2.1 above.

3.2.2.3 Vβ subfamily-specific semi-nested PCR

The optimised Vβ subfamily-specific semi-nested PCR protocol is detailed in section 2.4.4.2. cDNA (sections 2.4.1 - 2.4.3) from unfractionated PBMC was used undiluted

and cDNA prepared from T-cell clones was used diluted at either 1:5 or 1:10 in nuclease free water.

3.2.2.4 'Pan-Vβ' PCR

The optimised protocol for 'Pan-Vβ' PCR is detailed in section 2.4.4.2.

PCR products were subject to agarose gel electrophoresis, visualised and gel images analysed and stored as described in section 2.4.5.

3.2.3 Plasmid sub-cloning

(a) PCR amplified TCRβ chain transcripts

PCR products were purified using the Wizard PCR Preps DNA Purification System (section 2.4.6), sub-cloned into plasmid pGEM-T Easy vectors and used to transform High Efficiency Competent JM109 cells as described in section 2.4.7. Selected plasmid sub-clones were expanded overnight (section 2.4.7.2) and the presence of insert verified by restriction endonuclease analysis (section 2.4.7.1) of mini-preps of plasmid DNA (sections 2.4.7.3).

(b) Artefactual band in Vβ subfamily-specific semi-nested PCR

The artefactual band was extracted from agarose gel using the GeneClean II kit (Qbiogene, CA, USA) according to the manufacturer's instructions. Briefly, the DNA band was excised from an ethidium bromide stained agarose gel and transferred to a 1.5 ml centrifuge tube. The excised gel was weighed, and approximately 3 times the gel volume (assuming 1 g = 1 ml) of sodium iodide (NaI) was added to it; the gel/NaI was incubated at 45°C-55°C until the agarose dissolved. 10-15 µl of re-suspended Glassmilk was added to the gel/NaI and the mixture incubated at room temperature for 5-10 min with intermittent agitation every 1 min, before centrifugation at 10,000 *x g* for 5 min. The supernatant was discarded and the pellet washed in 30 x volume of New Wash a total of 3 times before being air-dried

at room temperature. The pellet was then re-suspended in 15 µl of nuclease-free water, centrifuged at 10,000 *x g* for 30 s and the eluted DNA recovered in the supernatant. The DNA was then sub-cloned into plasmid pGEM-T Easy vector and used to transform High Efficiency Competent JM109 cells as described above.

3.2.4 CDR3β-HDA

The protocol for CDR3β-HDA is given in section 2.4.8. When DNA from plasmid sub-clones was being utilised as the 'sample', then aliquots of 25ng DNA were used as the template for PCR amplification. Aliquots of 10 µl of H_2O were mixed with 'probes' prior to the de-naturing/re-annealing stage of CDR3β-HDA where indicated, to demonstrate that in the absence of other DNA the 'probes' did not generate heteroduplices.

For generation of FAM labelled PCR amplicons from sub-clone VB20.1 and subsequent CDR3β-HDA, the protocol for generation of Vβ20 'probe' was followed using VB20.1 DNA as template.

For the dilution series clones 592.15D ($V\beta2^+$) and (B) 485.144 ($V\beta13^+$) were serially diluted in unfractionated PBMC from animal 605 before RNA extraction and cDNA production (section 2.4.1-2.4.3).

3.2.5 *Theileria parva*-specific $CD8^+$ T-cell-lines and clones

Theileria parva-specific, *in vitro* $CD8^+$ T-cell enriched cell-lines were generated from animals 468 and 592, as described in section 2.3.3, and approximately 90 individual T-cell clones established from each them (section 2.3.4). The TCRβ chains expressed by these clones were sequenced following PCR amplification using the Vβ subfamily-specific semi-nested PCR (section 3.2.2.3) and the results are presented in Chapter 6 and Appendix L. cDNA from the cell-lines and the individual clones were subject to CDR3β-HDA as described above (section 3.2.5)

3.2.6 Sequencing

The sequences presented in this chapter were obtained from sequencing performed at either the Moredun Proteomics Facility (Moredun Research Institute, Penicuik, Scotland) or DBS genomics (Durham University, Durham, UK). Plasmid DNA and PCR products were prepared for sequencing using the Wizard Mini-preps DNA Purification System/alkali lysis method (section 2.4.7.3) and by the Wizard PCR Preps DNA Purification System (section 2.4.6) respectively and quantified (section 2.4.2). Plasmid sub-clone inserts were sequenced using either the T7 and/or SP6 specific sequencing primers and PCR products by the Cβ-specific 3'primer BCON-4 (Table 2.3).

3.3 Results

3.3.1 The known bovine TCRβV repertoire contains 46 genes in 17 subfamilies

Bovine TCRβV (Vβ) sequence data from studies using cDNA library analysis (Tanaka *et al.*, 1990), cross-species subfamily-specific PCR (Buitkamp *et al.*, 1993) and anchored PCR (Houston, 1997; Houston and Morrison 1999; Russell and MacHugh unpublished data) was compiled to form a database. Using the IMGT/V-QUEST software programme (Giudicelli *et al.*, 2004) the human Vβ gene to which each bovine Vβ gene shared highest nucleotide percentage identity was determined. From this analysis it was found that the gene previously designated as bovine Vβ12s1 actually had significantly higher nucleotide identity to human Vβ28s1 than to any of the Vβ12 members (80% compared to 66% at nucleotide level), and was therefore re-named as bovine Vβ28s1 (section 1.2.2). Within the previously unanalysed sequence data of Russell and MacHugh, a novel bovine subfamily, Vβ8, with two members (Vβ8s1 and Vβ8s2) was identified, as were two new members of the Vβ4 subfamily (Vβ4s3 and Vβ4s4). As noted previously, the two bovine Vβ17 members had less than 75% nucleotide homology to each other, but when compared to human sequences, both showed highest nucleotide identity to Vβ17 (Houston, 1997). The accumulated data gave a bovine repertoire of 46 Vβ genes distributed over 17 subfamilies, as summarised in Table 3.2.

3.3.2 Vβ-subfamily specific PCR amplification from cDNA

Using the sequence database, a set of Vβ subfamily-specific 5' primers were designed. For subfamilies with multiple Vβ gene sequences present in the database, the sequences were aligned using the ClustalW algorithm to identify conserved areas to which a primer could anneal; for some subfamilies (e.g.Vβ1) degeneracy had to be introduced towards the 5'end of the primer to accommodate all of the sequences. Due to lack of a conserved sequence of suitable length between Vβ17s1 and Vβ17s2 two primers, VB17a and VB17b were designed to be specific for Vβ17s2 and

Vβ17s1 respectively. Two 3' primers for the Cβ gene, BCint and BCext, were also designed using sequences in the GenBank database.

Vβ subfamily	Members identified by				Total number of members identified
	Tanaka *et al* (1990)	Buitkamp *et al* (1993)	Houston (1997), Houston and Morrison (1999)	Russell and MacHugh (unpublished)	
1	1		1,2,3,4,5,6,7,8*,9,10*,11*	4	11
2	1,3,6		1,2*,3,4,5*,6,7*	3,5*,6,7*	7
3	1		1		1
4	1		2*	3*,4*	4
6	1	2	3*	2	3
7	1		1,2*	1	2
8				1*,2*	2
9			1*		1
10			1,2,3		3
13			1,2,3	1	3
14			1		1
15	1,2		1,2		2
16			1		1
17	1		2*	1,2*	2
20			1	1	1
24			1	1	1
28	1		1		1
Total	12	1	39	16	46

Table 3.2 The known bovine Vβ repertoire. The members of each Vβ subfamily (column 1) identified by Tanaka *et al.* (1990) - column 2, Buitkamp *et al.* (1993) - column 3, Houston (1997) and Houston and Morrison (Houston and Morrison, 1999) - column 4 and Russell and MacHugh, unpublished - column 5, are shown. Genes suffixed by * have not be deposited in GenBank and the sequences are derived from either Houston (1997) or unpublished data provided by Russell and MacHugh (section 2.5.2). These sequences are detailed in Appendix E. The total number of known members belonging to each Vβ subfamily is shown in column 6.

The 5' Vβ subfamily-specific primers were used in separate reactions, with the 3' Cβ-specific primer BCext, to amplify cDNA from un-fractionated bovine PBMC; primers that failed to generate product of the anticipated size were re-designed. The

product from each reaction was sub-cloned into pGEM-T plasmid vectors and the sequences for two plasmid sub-clones from each reaction were determined (Table 3.3). All belonged to the anticipated Vβ subfamily. However, the Vβ gene sequence in sub-clones VB1.1, VB10.2, VB13.1 and VB13.3 were sufficiently different from previously identified members of the respective subfamilies to be considered new members - Vβ1s12, Vβ10s4, Vβ13s4 and Vβ13s5 respectively (section 2.5.1). Descriptions of these genes are given in Chapter 4. Although the sequence of the Vβ region of sub-clone VB20.1 was identifiable as Vβ20s1, it was not possible to obtain sequence for the CDR3 or Jβ gene regions; later work demonstrated that this was because VB20.1 was not mono-clonal (section 3.3.4).

Plasmid sub-clone	Sequence				
	Vβ gene	FR	CDR3	FR	Jβ gene
1.4	1s1	TGTCAAAGC C A S	AGCCGCGCTTCCAATTATGACTAT S R A S N Y D Y	CACTTCGGC H F G	1s2
1.1	1s12	TGTGCCAGC C A S	AGCCAAGGGCAATCCGAGACGCTG S Q G Q S E T L	TACTTCGGG Y F G	2s4
2.4	2s6	TGCAGTGCT C S A	CCAACCCGGGACTGGGTGTATGAGCAG P T R D W V Y E Q	TATTTCGGC Y F G	1s2
2.2	2s3	TGCAGTGCT C S A	GAACGCAACAACCCTCTGTAT E R N N P L Y	TTTGAAGGG F E G	3s3
3.3	3s1	TGCGCCAGC C A S	AGTATGGGTTCCGGGACAGCTGGAGCGATCCAG S M G S G T A G A I Q	TACTTCGGG Y F G	3s5
3.1	3s1	TGCGCCAGC C A S	AGTGACGGACAGCAAAGCACTCAG M D G Q Q S T Q	TACTTCGGC Y F G	2s3
4.3	4s3	TGCAGTGAT C S D	CGGACAGCAGCTTCTGGAGACTAT R T A A S G D Y	CACTTCGGC H F G	1s2
4.1	4s3	TGCAGTGCT C S A	ATTCGACAGCACAGCAGCAACCAGGCACAG I R Q H S S N Q A Q	AACTTTGGA N F G	1s4
6.1	6s2	TGTGCCAGC C A S	AGCCCAGGACTTCGGGGGGGTATGAACACCCAGCCCCTG S P G L R G G M N T Q P L	TACTTTGGA Y F G	3s2
6.2	6s1	TGTGCCAGC C A S	AGCCCGGGGGGTGGGGGGGGCACTCAGATCCGG P G G G G G T Q I R	TACTTCGGG Y F G	3s5
7.3	7s2	TGCGCCAGC C A S	AGTCGAGGGGGGAACGTGGAGCAG S R G G N V E Q	TATTTCGGC Y F G	3s7
7.2	7s1	TGCGCCAGC C A S	AGTAGAGATCGGAGGGGGGGCAACAACCCTCTG M R D R R G G N N P L	TATTTTGGA Y F G	3s3
8.2	8s1	TGTGCCACC C A T	TGGGGGCCCGACCAAAAACATTCCTATGAGCAG G P D Q K H S Y E Q	TATTTCGGC Y F G	3s7

Table 3.3 Sequences of TCRβ chain inserts of the panel of plasmid sub-clones generated.

Plasmid sub-clone	Sequence				
	Vβ gene	FR	CDR3	FR	Jβ gene
8.1	8s1	TGTGCCAGC C A S	AGCAAAGATCTCCTCAGCAAT S K D L L S N	CACTTCGGC H F G	1s2
9.4	9s1	TGTGCCAGC C A S	AGCCTCTCAGACACGCAG S L S D T Q	TACTTCGGC Y F G	3s4
9.1	9s1	TGTGCCAGC C A S	GACCGGGGGGCGGCAGGTGGCCCTCTG D R G A A G G P L	TATTTTCCA Y F G	3s3
10.2	10s4	TGTGCCAGT C A S	AGCCGAGTCGGGGGGGTGGACAACCCTCTG S R V G G V D N P L	TATTTTGGA Y F G	3s3
10.1	10s4	TGTGCCAGT C A S	AACCAAGATCGGACTTCAATACGTGGGGAGCTG N Q D R T S I R G E L	CACTTCGGG H F G	3s1
13.1	13s4	TGCGCCAGC C A S	AGTCTGGGGGGCGATGAGGAGCAGCAC S L G G D E E Q H	CACTTCGGG H F G	2s1
13.3	13s5	TGTACCAGC C T S	AGCCTAGGACGGCTCAACAACCCTCTG S L G R L N N P L	TATTTTGGA Y F G	3s3
14.6	14s1	TGTGCCAGC C A S	AGTGAGGGGACAAGGGGCTTTGCAGGCGCCGCCCTG S E G T R G F A G A A L	ACCTTCGGG T F G	3s6
14.7	14s1	TGTGCCAGC C A S	AAGGGACCTCAGAACAGCCCTCTG K G P Q N S P L	TATTTTGGA Y F G	3s3
15.3	15s2	TGTGCCAGC C A S	GCAAGGGAGCTAGACTAT A R E L D Y	CACTTGGGC H F G	1s2
15.1	15s2	TGTGCCAGC C A S	ACCAGGGGACAGCAGGGCGCCAGCACTCAG T R G Q Q G A S T Q	TACTTCGGC Y F G	2s3
16.5	16s1	TGTGCCAGC C A S	AGCCGAGGTTTCGAGGAGCAG S R G F E E Q	CACTTCGGG H F G	2s1
16.1	16s1	TGTGCCAGC C A S	AGCCGAGGCGGGACAGCCGGGACTCAGATCCAG S R G G T A G T Q I Q	TACTTCGGG Y F G	3s5
17A.7	17s2	TGCTCTGGG C S G	GGGACAGCAGGCTCTACTCAGATCCAG G T A G S T Q I Q	TACTTCGGG Y F G	3s5
17A.6	17s2	TGCTCTGGG C S G	GGAAAGACAGCCCTTAATGACTAT G K T A L N D Y	CACTTCGGC H F G	1s2
17B.5	17s1	TGTGCTGCC C A A	ATGGGGGGTGGAGGCCAAAGCACTCAG M G G G G Q S T Q	TACTTCGGC Y F G	2s3
17B.1	17s1	TGTGCTGCT C A A	TCGACTTCGGGGGCCTATGAGCAG S T S G A Y E Q	TATTTCGGC Y F G	3s7
20.4	20s1	TGTGCCTGG C A W	AGTCGCGGGGACCTTCGAATAAGCAACAACCCTCTG S R G D L R I S N N P L	TATTTTGGA Y F G	3s3
20.1	20s1	Not able to sequence			
24.5	24s1	TGTGCCAGC C A S	AGCAGAGATCGGACTTCGATCACAGACACTCAG S R D R T S I T D T Q	TACTTCGGC Y F G	3s4
24.1	24s1	TGTGCCAGC C A S	AGCAGGTGGACAGCAGATAGCACAGACACGCAG S R W T A D S T D T Q	TACTTCGGC Y F G	3s4
28.4	28s1	TGTGCCAGC C A S	GCT GCC GGG ACA GTT GAC TAT A A G T V D Y	CACTTCGGC H F G	1s2
28.3	28s1	TGCGCCAGC C A S	GCT GCC CCG GCA GGG GCC GGA TAT GAG CAG A A P A G A G Y E Q	TATTTCGGC Y F G	3s7

Table 3.3 (continued) Sequences of TCRβ chain inserts of the panel of plasmid sub-clones generated.

The specificity of the Vβ subfamily-specific primers was assessed using diluted DNA from a panel of the plasmid sub-clones. Several of the primers displayed marked cross-reactivity, amplifying genes belonging to various Vβ-subfamilies and required re-design (data not shown). Ultimately, a set of primers, all but 3 of which only amplified Vβ genes of the relevant subfamily when tested against the panel of plasmid sub-clones (i.e. were subfamily-specific - e.g. VBext13 - Figure 3.1a), was acquired. Primers VB2, 10 and 16 non-specifically amplified Vβ genes of other subfamilies (e.g. VBext2 - Figure 3.1b) as detailed in Table 3.4.

To achieve improved subfamily specificity it was decided to design an additional set of Vβ subfamily-specific primers to be used in conjunction with the initial set (henceforth referred to as VBext primers) in semi-nested PCR (snPCR). As with the VBext primers, several of these VBint primers had to be re-designed because when used in single-round PCR (i.e. non-nested) they either failed to generate product from cDNA of unfractionated PBMC or displayed cross-reactivity when examined for Vβ subfamily-specificity against the panel of plasmid sub-clones. The final set of VBint primers was subfamily-specific, with the exception of VB1int, VB13int and VB14int which amplified Vβ genes of other subfamilies (Table 3.4).

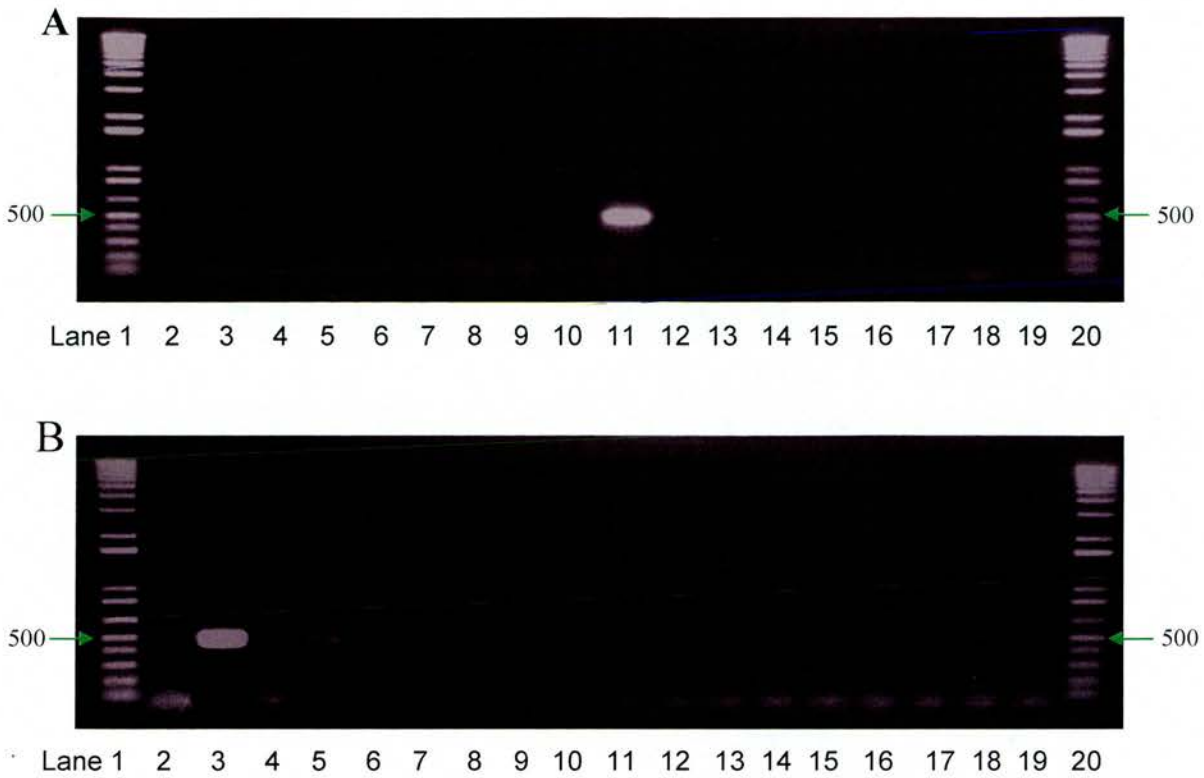
Vβ subfamily-specific primer	**Cross-reactive amplification of plasmid sub-clone**	**Vβ gene contained in plasmid sub-clone**
VBext2	4.1	4s3
VBext10	13.1	13s4
VBext16	7.2	7s1
VBint1	6.2	6s1
	13.1	13s4
VBint13	24.1	24s1
VBint14	13.1	13s4
	17A.6	17s2

Table 3.4 Vβ subfamily-specific primers demonstrating cross-reactive amplification of plasmid sub-clones containing Vβ genes from other subfamilies.

Importantly, the pattern of cross-reactivity displayed by the two sets of 5' primers suggested that when used in combination in snPCR, amplification should be Vβ subfamily-specific (i.e. none of the VBext/VBint primer pairs cross-reactively amplified the same Vβ subfamily). Indeed, the results indicated that replacement of VBext2, 10 and 16 with their VBint counterparts would have sufficed in creating a Vβ subfamily-specific set of 5' primers. Nevertheless it was decided to develop a snPCR protocol as the combined use the two 5'primers in snPCR would; (i) afford more assured Vβ subfamily-specificity of PCR amplification and (ii) if required could provide greater sensitivity without the compromise in specificity frequently seen if single primer sets are used in a high number of PCR cycles.

A snPCR protocol was established that allowed the generation of a single band of the anticipated size by each pair of VBext/VBint subfamily-specific 5'primers from cDNA of un-fractionated PBMC (Figure3.2a). Using this protocol cDNAs from 22 T-cell clones that had been generated within the laboratory were amplified. The majority of cDNAs were amplified by only one set of Vβ subfamily-specific primers (e.g. 485.280 - Figure 3.2b), but some were amplified by 2 or more sets (e.g. 485.71 - Figure3.2c). Direct sequencing of the products confirmed that in each case amplification was Vβ subfamily-specific (Table 3.5) and that some of the T-cells expressed 2 or more TCRβ chains, presumably because they were not mono-clonal, expressed a non-functional as well as a functional TCRβ chain or lacked allelic exclusion (Davodeau *et al.*, 1995; Padovan *et al.*, 1995).

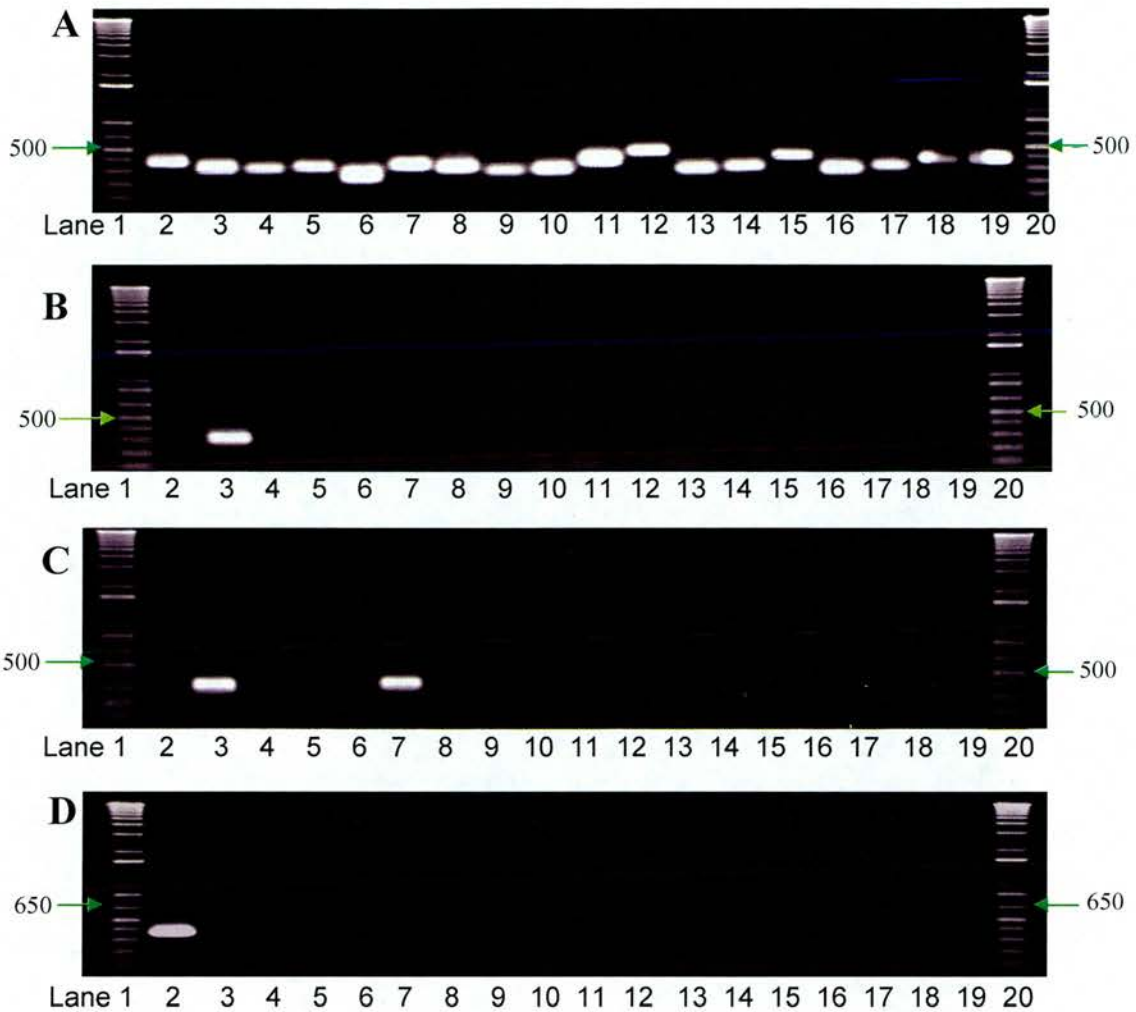
Within the 24 unique TCRβ chains expressed by the 22 clones, 18 of the 50 previously sequenced Vβ genes and 4 novel Vβ genes (designated as Vβ2s8, Vβ4s5, Vβ10s5 and VB17s3 - described in Chapter 4) were identified. These Vβ genes were distributed over 14 of the 17 bovine Vβ subfamilies (no Vβ8, Vβ15, Vβ17s2 or Vβ28 genes were identified), demonstrating the effectiveness of the snPCR over a wide coverage of the Vβ repertoire.



Lane

1 - Molecular markers	6 - plasmid VB6.2	11 - plasmid VB13.1	16 - plasmid VB17B.1
2 - plasmid VB1.1	7 - plasmid VB7.2	12 - plasmid VB14.6	17 - plasmid VB20.1
3 - plasmid VB2.2	8 - plasmid VB8.1	13 - plasmid VB15.1	18 - plasmid VB24.1
4 - plasmid VB3.1	9 - plasmid VB9.1	14 - plasmid VB16.1	19 - plasmid VB28.3
5 - plasmid VB4.1	10 - plasmid VB10.1	15 - plasmid VB17A.6	20 - Molecular markers

Figure 3.1 Specificity of Vβ subfamily-specific primers. PCR amplification of DNA from a panel of plasmid sub-clones by (A) VBext13 and (B) VBext2. VBext13 is subfamily-specific amplifying only DNA from plasmid sub-clone VB13.1 (lane 11) which expresses the Vβ13s4 gene. VBext2 amplifies sub-clone VB4.1 (lane 5) which expresses the Vβ4s3 gene as well as sub-clone VB2.2 (lane 3) which expresses the Vβ2s3 and therefore is not subfamily-specific.



Lane

1-Molecular markers	6-VB6*2ext/VB6*1int	11-VB13*1ext/VB13*1int	16-17B*1ext/VB17B*1int
2-VB1*1ext/VB1*1int	7-VB7*2ext/VB7*1int	12-VB14*2ext/VB14*1int	17-VB20*1ext/VB20*1int
3-VB2*1ext/VB2*1int	8-VB8*2ext/VB8*1int	13-VB15*2ext/VB15*2int	18-VB24*3ext/VB24*1int
4-VB3*3ext/VB3*1int	9-VB9*2ext/VB9*1int	14-VB16*5ext/VB16*2int	19-VB28*2ext/VB28*1int
5-VB4*2ext/VB4*1in	10-VB10*2ext/VB10*1int	15-VB17A*1ext/VB17A*3int	20-Molecular markers

Figure 3.2 PCR amplification using the Vβ subfamily-specific snPCR protocol. (A) Amplification of cDNA from unfractionated PBMC - showing generation of product of the anticipated size for each subfamily (lanes 2-19). (B) Amplification by VBext2/VBint2 (lane 3) of cDNA from T-cell clone 485.280 which expresses a $V\beta2^+$ TCRβ chain - Table 3.5. (C) Amplification by VBext2/VBint2 and VBext7/VBint7 (lanes 3 and 5) of cDNA from T-cell clone 485.71 which expresses $V\beta2^+$ and $V\beta7^+$ TCRβ chains - Table 3.5. (D) Artefactual band of ~650bp generated in reactions producing no Vβ subfamily-specific product (lanes 3-19) in amplification of cDNA from the $V\beta1^+$ clone 592.10D.

T-cell clone	Sequence				
	Vβ gene	FR	CDR3	FR	Jβ gene
485.53	2s3	TGCAGTGCC C S A	TTCAAGTTCGGGGGGCCAGACACAGACACGCAG F K F G G P D T D T Q	TACTTCGGC Y F G	3s4
485.71	2s3	TGCAGTGCC C S A	TTCAAGTTCGGGGGGCCAGACACAGACACGCAG F K F G G P D T D T Q	TACTTCGGC Y F G	3s4
	7s1	TGCGCCAAC C A N	AGATCGAAACTGGGGGATGATGAGGAGCTG R S K L G D D E E L	CACTTCGGG H F G	3s1
485.130	2s3	TGCAGTGCC C S A	TTCAAGTTCGGGGGGCCAGACACAGACACGCAG F K F G G P D T D T Q	TACTTCGGC Y F G	3s4
485.182	2s3	TGCAGTGCC C S A	TTCAAGTTCGGGGGGCCAGACACAGACACGCAG F K F G G P D T D T Q	TACTTCGGC Y F G	3s4
485.280	2s3	TGCAGTGCC C S A	TTCAAGTTCGGGGGGCCAGACACAGACACGCAG F K F G G P D T D T Q	TACTTCGGC Y F G	3s4
485.135	17s3	TGTGCTGCT C A A	AGTAGAGATCGAGGCTCCGGTGAGCGG S R D R G S G E R	TATTTTGGC Y F G	2s5
	13s3	TGTGCCAGC C A S	GGTATGGCAGCTAGCCAAAGCACTCAG G M A A S Q S T G	TACTTCGGC Y F G	2s3
485.198	10s5	TGTGCCAGC C A S	AGCCGCTCATACAGCAGTGAGCGG S R S Y S S E R	TATTTTGGC Y F G	2s5
	13s1	TGTGCCAGC C A S	AGTCATGGACTTCGGGGGGGGGACAACCCTCTG S H G L R G G D N P L	TATTTTGGA Y F G	3s3
485.144	13s1	TGTGCCAGC C A S	AGTCATGGACTTCGGGGGGGGGACAACCCTCTG S H G L R G G D N P L	TATTTTGGA Y F G	3s3
592.10D	1s7	TGTGCCAGC C A S	AGCCTCGGGGTGGGGGGTGGTGGCTCAGAGACGCTG S L G V G G G G S E T L	TACTTCGGG Y F G	2s4
592.14D	14s1	TGTGCCAGC C A S	AGTGAATCGCAAGGGGCGAATTATGACTAT S E S Q G A N Y D Y	CACTTCGGC Y F G	1s2
592.15D	2s4	TGCAGTGCT C S A	CAATGGGGGGGTTCCTATGCGGCGCAG Q W G G S Y E E Q	CACTTCGGG H F G	2s1
592.36D	1s7	TGTGCCAGC C A S	AGCCAAGATCTGACAGCTGGATATGGGGAGCTG S Q D L T A G Y G E L	CACTTCGGG H F G	3s1
633.4	16s1	TGTGCCAGC C A S	AGCCGTGGCGGGCTAGGCACTGAGGTT S R G G L G T E V	TTCTTTGGA F F G	1s1
633.49	3s1	TGCGCCAGC C A S	AGTTCATATGGAAGCAACAACCCTCTG S S Y G S N N P L	TATTTTGGA Y F G	3s3
	24s1	TGTGCCAGC C A S	AGCATGGGGGGTGGATGGTCGGAGACGCTG S M G G G W S E T L	TACTTCGGG Y F G	2s4
663.6	1s5	TGTGCCAGC C A S	AGCGATAGAGACTGGGGGGTGAATTCCTATGAGCAG S D R D W G V N S Y E Q	TATTTCGGC Y F G	3s7
	4s3	TGCAGTGCT C S A	GGCCGCGGGGACCGATTATATGACTAT G R G D R L Y D Y	CACTTCGGC H F G	1s2
	6s2*	TGTGCCACC C A T	GGGAGGGGGCACTCAG	TACTTCGGC Y F G	2s3
	20s1	TGTGCCTGC C A C	TCTCGGGGCAACAACCCTCTG S R G N N P L	TATTTTGGA Y F G	3s3

Table 3.5 Sequences of TCRβ chains expressed by 22 T-cell clones obtained using the Vβ subfamily-specific snPCR protocol.

T-cell clone	Sequence				
	Vβ gene	FR	CDR3	FR	Jβ gene
663.63	4s5	TGCAGTGCT C S A	GGAGAGCRAGGCGTGAGGAACCCTCTG G E L G V R N P L	TATTTTCCA Y F G	3s3
663.75	9s1	TGTGCCAGC C A S	AGCCCAACCTCTCGAGGGCCCAACAACCCTCTG S P T S R G P N N P L	TATTTTGGA Y F G	3s3
663.139	9s1	TGTGCCAGC C A S	AGGCACAGCAGGGATGACTAT R H S R D D Y	CACTTCGGC H F G	1s2
T28.27	2s8	TGTGGTGCT C G A	AGAGTCGGCTATGAGCAG R V G Y E Q	TATTTCGGC Y F G	3s7
F104	24s1	TGTGCCAGC C A S	CGTGGCGAGAACGCCGCGCAGCTG R G E N A A Q L	TACTTTGGA Y F G	2s2
T28.1	2s5	TGTGGTGCG C G A	GAGGGGGGACAGTCCTATGAGCAG E G G Q S Y E Q	TATTTCGGC Y F G	3s7
7193TpM	2s7	TGTGGTGCT C G A	AGGCAGTCGGGGGGACTGAACACCCAGCCCCTG R Q S G G L N T Q P L	TACTTTGGA Y F G	3s2

Table 3.5 (continued) Sequences of TCRβ chains expressed by 22 T-cell clones obtained using the Vβ subfamily-specific snPCR protocol. * Sequence is out of frame and therefore non-functional.

The sensitivity of the snPCR protocol has not been fully assessed for all of the primer pairs. However, using Vβ3-specific primers with cDNA derived from a $V\beta3^+$ T-cell clone (633.49), it was demonstrated that product could be generated from the equivalent of ~300 cells (data not shown). This level of sensitivity was considered sufficient for the purpose of identifying the TCRβ chains expressed by individual T-cell clones.

A frequent artefact of the snPCR protocol was the generation of a product of approximately 650bp in reactions that generated no Vβ subfamily-specific product (Figure 3.2d). This band was extracted from one gel, purified and sub-cloned into a plasmid vector for sequencing. The resulting sequence was found to have 99% nucleotide identity with a sequence in the bovine genome database (locus XM_595623 on scaffold NW_932240.1/Bt9_WGA1710_2 - Appendix F), which automated annotation has predicted to encode a homologue of human nesprin-1. The artefact displayed 87% nucleotide sequence identity with the deposited sequence of human nesprin-1 protein (GenBank accession number AF495910 - Zhang *et al.*, 2002 - Appendix B). Analysis of the sequence revealed an annealing site for BCint and the reverse complement of BCint at opposing ends of the product (Appendix F). As the

size of the artefact was significantly greater than any of the anticipated Vβ-subfamily specific products and the artefact band was generally weak, it was considered unlikely to interfere in subsequent application of the snPCR and therefore no efforts were made to eliminate it by adjusting the protocol.

3.3.3 Design of a 'Pan-Vβ' primer

None of the Vβ subfamily-specific primers were able to amplify cDNA from a cloned, *Theileria parva*-transformed cell line (468TpM), which was shown by FACS analysis to express CD3 but not the γ/δ TCR and was therefore considered to be an αβ T-cell (data not shown). This suggested that 468TpM expressed a TCRβ chain containing a Vβ gene that was either (i) from a novel bovine Vβ subfamily or (ii) a novel member of a known subfamily that lacked the conserved sequence against which one or both of the subfamily-specific primers annealed.

Degenerate primers directed against sequences conserved between V subfamilies have been used in rodents (Broeren *et al.*, 1991), humans (Obata *et al.*, 1993), dogs (Ito *et al.*, 1993) and pigs (Baron *et al.*, 2001) to amplify TCR chains in which the sequence of the V gene is unknown. It was decided to use this methodology to try to isolate the TCRβ chain of 468TpM.

Alignment of all the sequenced bovine Vβ genes using the ClustalW algorithm revealed that the only universally conserved sequence was the 5 base TGGTA that includes the codon for the conserved tryptophan at amino acid position 41 of the Vβ domain. Therefore, a bovine 'Pan-Vβ' 5' primer was designed with these 5 bases at the 3'end. An additional 9 bases of the primer contained various degrees of degeneracy based on the known bovine Vβ gene sequences (Table 3.6). Four guanines were added to the 5'end of the primer to aid stabilization of annealing.

Base	14	13	12	11	10	9	8	7	6	5	4	3	2	1
Primer sequence	N	N	N	V	T	D	T	H	Y	T	G	G	T	A
A	12	10	3	26		5	1	36						54
C	2	33	16	7	1		1	6	48					
G	10	7	4	21		39					54	54		
T	30	4	31		53	10	52	12	6	54			54	

Table 3.6 Design of the bovine 'Pan-Vβ' 5' primer. All of the sequenced bovine Vβ gene sequences were aligned by the ClustalW algorithm and the frequency of all four nucleotides (A,C,G and T) at the nine positions preceeding the conserved TGGTA at positions 100-104 (IMGT Unique Numbering System) in the known Vβ genes was determined. Based on this data the sequence of the 14 3' bases of the primer were established with degeneracy introduced as appropriate (Y=C/T; N=A/C/G/T; V=A/C/G; D=A/G/T; H=A/C/T).

Used in conjunction with the two Cβ-specific 3'end primers (PVBrev1 and PVBrev2 - Table 2.3) in a snPCR protocol the 'Pan-Vβ' primer was able to amplify cDNA from 468TpM, which expressed a Vβ gene subsequently designated as a member of a novel bovine subfamily, Vβ5 (described in Chapter 4).

3.3.4 A bovine CDR3β heteroduplex analysis (CDR3β-HDA) method

An overview of the basic principle of the CDR3β-HDA method developed in this study is given in section 2.4.8.1 and details of the method are provided in section 2.4.8.2.

The requirement for 5'end (Vβ) and 3' (Cβ) end complementarity between TCRβ DNA strands forming a heteroduplex dictates that the CDR3β-HDA method has a substantial degree of inherent Vβ-subfamily specificity. Using the non-specifically amplified products of VB13.1 and VB17A.6 generated by the VB14int 5'primer (Figure 3.3a), it was demonstrated that this was indeed the case. Whilst a heteroduplex could be formed between the Vβ14 'probe' (generated from VB14.6) and the VB14int/BCext amplified product of VB14.7 (Figure 3.3b lane 2), neither of the products amplified from VB13.1 or VB17A.6 was able to form a heteroduplex

with the Vβ14 'probe' (Figure 3.3b lanes 3 and 4). Similarly, cross-reactive products generated by other Vβ-subfamily-specific primers could not form heteroduplices with the relevant Vβ subfamily probe (data not shown). Consequently it was decided to use only a single round PCR to generate amplicons used in the CDR3β-HDA method, as sufficient Vβ-subfamily specificity would be present in the system.

The ability to form CDR3β heteroduplices for each Vβ subfamily was verified using products of the respective subfamily but derived from different plasmid sub-clones than those selected for 'probe' generation (examples given in Figure 3.4a). Importantly, for the majority of the multi-membered subfamilies the sub-clones contained Vβ genes that were different from those present in the 'probes', demonstrating the ability to form heteroduplices between different members of the same Vβ subfamily (e.g. Vβ2 - the Vβ2 'probe' is generated from plasmid sub-clone VB2.2 which contains Vβ2s3 and the plasmid sub-clone VB2.4 contains Vβ2s6).

Amplicons from plasmid sub-clone VB20.1 formed two heteroduplex bands with the Vβ20 'probe' (Figure 3.4b lane 2) and furthermore were capable of forming heteroduplices in the absence of the 'probe' (Figure 3.4b lane 3). This infers that sub-clone VB20.1 contains 2 clones with dissimilar CDR3β regions and explains why it was not possible to obtain a sequence for the CDR3β/Jβ regions of this plasmid (section 3.3.2).

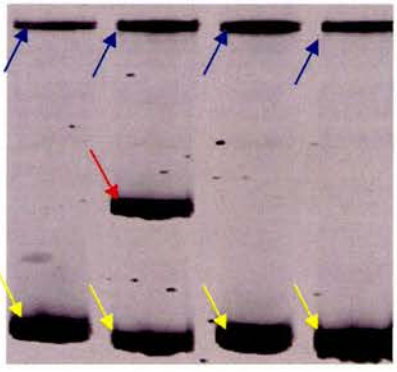
A

Lane 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Lane

1 - Molecular markers	6 - plasmid VB6.2	11 - plasmid VB13.1	16 - plasmid VB17B.1
2 - plasmid VB1.1	7 - plasmid VB7.2	12 - plasmid VB14.7	17 - plasmid VB20.1
3 - plasmid VB2.2	8 - plasmid VB8.1	13 - plasmid VB15.1	18 - plasmid VB24.1
4 - plasmid VB3.1	9 - plasmid VB9.1	14 - plasmid VB16.1	19 - plasmid VB28.3
5 - plasmid VB4.1	10 - plasmid VB10.1	15 - plasmid VB17A.6	20 - Molecular markers

B



Lane 1 - Vβ14 probe + H_2O

2 - Vβ14 probe + PCR product from plasmid sub-clone VB14.7

3 - Vβ14 probe + PCR product from plasmid sub-clone 13.1

4 - Vβ14 probe + PCR product from plasmid sub-clone 17A.6

Figure 3.3 Vβ subfamily-specificity of CDR3β-HDA. (A) PCR amplification of DNA from a panel of plasmid sub-clones with VB14int/BCext is not subfamily specific - as well as specific amplification of sub-clone VB14.7 (Vβ14$^+$ - lane 12), sub-clones VB13.1 (Vβ 13$^+$ - lane 11) and VB 17A.6 (Vβ17$^+$ - lane 15) are cross-reactively amplified. (B) CDR3β-HDA has inherent Vβ subfamily specificity as the Vβ14 'probe' can form a heteroduplex with the amplified product from sub-clone VB14.7 (lane 2) but not the amplified products from sub-clones VB13.1 (lane 3) or VB17A.6 (lane 4). Heteroduplex (red arrows), homoduplices (yellow arrows) and single-stranded 'probe' DNA (blue arrows) have been indicated.

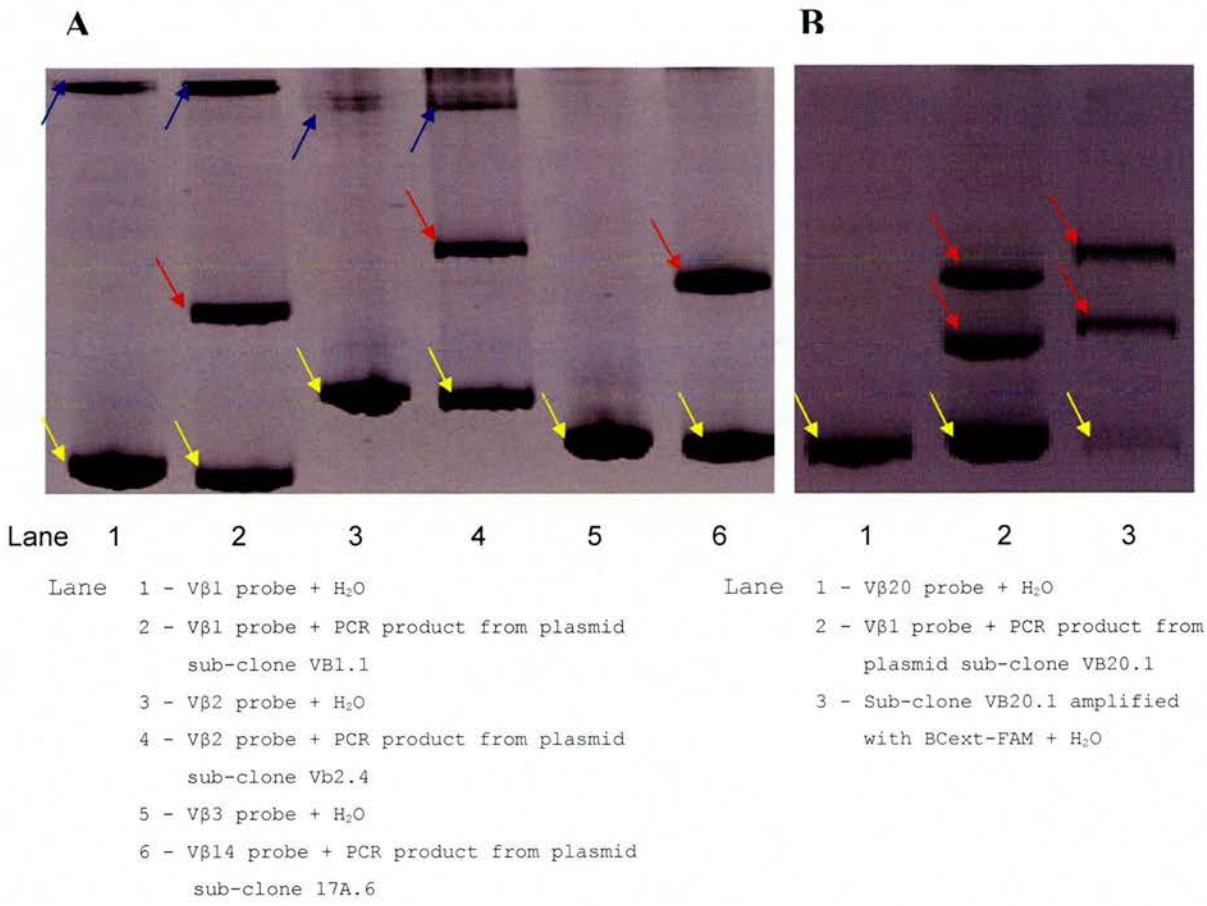


Figure 3.4 CDR3β-HDA formation. (A) Examples of heteroduplex formation between the; Vβ1 'probe'/amplified PCR product from plasmid sub-clone VB1.1 (lane 2a), Vβ2 'probe'/amplified PCR product from plasmid sub-clone VB2.4 (lane 4a) and Vβ3 'probe'/amplified PCR product from plasmid subclone VB3.3 (lane 6a). When run in the absence of other DNA none of the 'probes' generate heteroduplex bands (lanes 1a, 3a, 5a). (B) Vβ20 subfamily heteroduplex formation. Two heteroduplices are formed between the Vβ20 'probe'/amplified PCR product from plasmid sub-clone VB20.1 (lane 2b). When run in the absence of other DNA the Vβ20 'probe' only formed a homoduplex (lane 1b) whereas if sub-clone VB20.1 was amplified with the labelled BCext-FAM 3'primer and then subjected to CDR3β-HDA in the absence of any other DNA, two heteroduplices were formed, indicating that sub-clone VB20.1 was not composed of mono-clonal TCRβ. This would explain the inability to sequence the CDR3 sequence of the TCRβ chain insert of this plasmid (section 3.3.2). Heteroduplices (red arrows), homoduplices (yellow arrows) and single-stranded 'probe' DNA (blue arrows) have been indicated.

3.3.4.1 Clonal specificity and sensitivity of clonal detection by CDR3β-HDA

A key feature of CDR3β-HDA is that each expanded T-cell clonotype is represented as a distinct and reproducible band that constitutes a unique 'molecular footprint' (Maini *et al.*, 1998). This enables (i) different expanded clones to be distinguished from each other and (ii) clones to be re-identified in multiple samples - either from different time-points in longitudinal studies or from different T-cell populations (e.g. comparisons between *in vitro* and *ex vivo* or local and systemic T-cell populations). Using a small panel of Vβ2$^+$ clones/sub-clones it was possible to establish that T-cells expressing identical TCRβ chains (i.e. of the same clonotype) had a common CDR3β-HDA 'molecular footprint', whilst those expressing TCRβ chains with the same Vβ subfamily but different CDR3β regions had dissimilar CDR3β-HDA migration patterns (Figure 3.5).

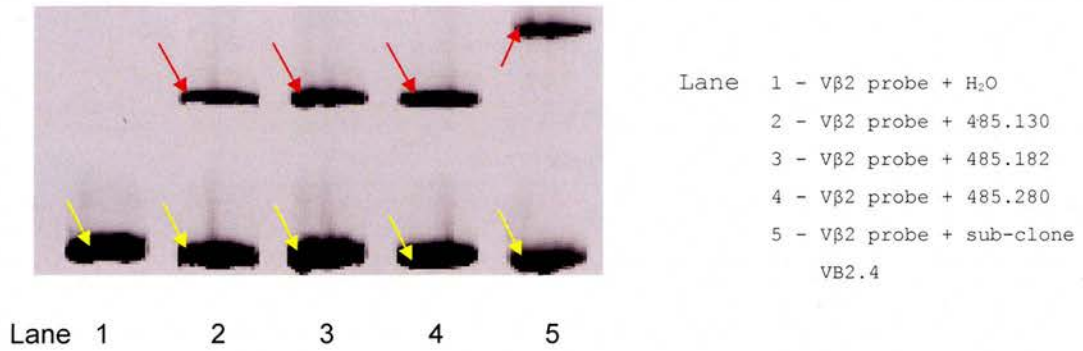
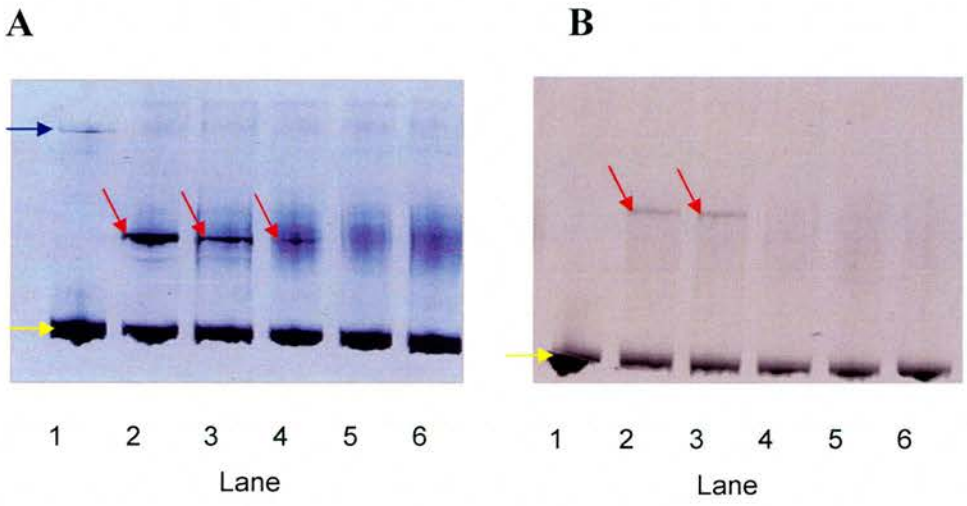


Figure 3.5 CDR3β-HDA migration patterns of Vβ2$^+$ bovine TCRβ chains. Vβ2$^+$ T-cell clones 485.130 (lane 2), 485.182 (lane 3) and 485.280 (lane 4) express identical TCRβ chains (Table 3.5) and have a common CDR3β-HDA 'molecular footprint' in CDR3β-HDA. The CDR3β-HDA migration pattern of the Vβ2$^+$ TCRβ chain insert of plasmid sub-clone VB2.4 (lane 5) is dissimilar to that of the three 485 T-cell clones (Table 3.3), reflecting its expression of a different CDR3β sequence. Heteroduplices (red arrows) and homoduplices (yellow arrows) have been indicated.

By seeding Vβ2$^+$ and Vβ13$^+$ *in vitro* derived T-cell clones independently into unfractionated PBMC it was possible to evaluate the sensitivity of the CDR3β-HDA method. The sensitivity differed between the two clones; a distinct band representing

the Vβ2$^+$ clone was clearly visible at a dilution of 1:10,000 (Figure 3.6a), but the Vβ13$^+$ clone heteroduplex band was only evident down to a dilution of 1:1000 (Figure 3.6b).



Lane 1 - Vβ probe + H_2O
2 - Vβ probe + clone at 1:100 in unfractionated PBMC
3 - Vβ probe + clone at 1:1,000 in unfractionated PBMC
4 - Vβ probe + clone at 1:10,000 in unfractionated PBMC
5 - Vβ probe + clone at 1:100,000 in unfractionated PBMC
6 - Vβ probe + clone at 1:1,000,000 in unfractionated PBMC

Figure 3.6 Sensitivity of detection of T-cell clonal expansions by CDR3β-HDA. Serial ten-fold dilution of T-cell clones (A) 592.15D (Vβ2$^+$) and (B) 485.144 (Vβ13$^+$) in unfractionated PBMC. A distinct heteroduplex band (red arrow) representing the TCRβ chain expressed by 592.15D is clearly visible in dilutions down to 1:10,000 (lane 4a). The heteroduplex band representing the TCRβ chain expressed by 485.144 (red arrow) is only evident at dilutions down to 1:1,000 (lane 3b). In the V β2 series it is notable that a polyclonal smear (section 2.4.8.1) becomes more prominent as the seeded clones are diluted out (lanes 2a-6a). The homoduplices formed by the labelled 'probes' are indicated by yellow arrows and single-stranded labelled 'probe' DNA by blue arrows.

3.3.4.2 Analyzing a complex T-cell population by CDRβ3 -HDA

In humans CDRβ3 -HDA techniques have been used principally to provide a rapid, global analysis of the clonal composition of complex T-cell populations (Lopes *et al.*, 2003; Maini *et al.*, 2000). The bovine CDRβ3-HDA method was used to examine the clonal composition of *T. parva*-specific *in vitro* $CD8^+$ T-cell enriched cell-lines derived from the PBMC of two immunised animals (Figure 3.7a and b). CDRβ3 - HDA revealed that both cell-lines were highly oligoclonal, with many distinct heteroduplex bands distributed amongst most of the Vβ subfamilies.

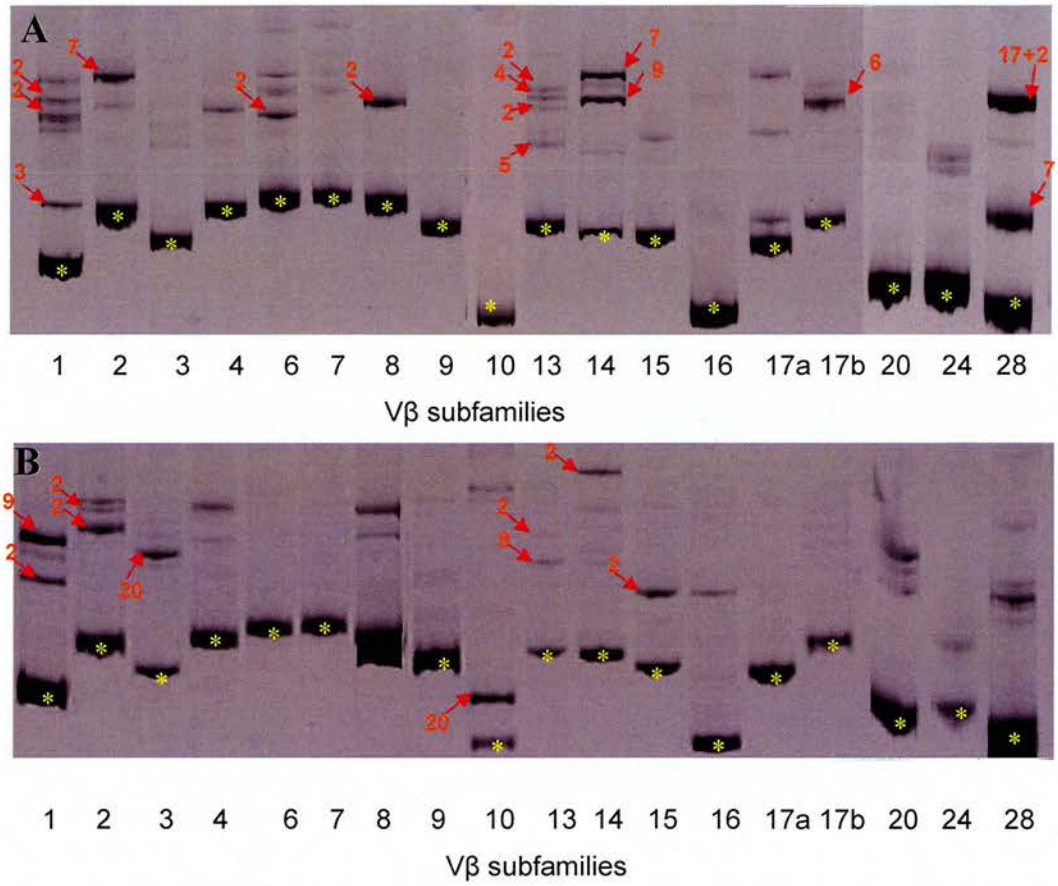


Figure 3.7 CDR3β-HDA of complex bovine T-cell populations. Analysis of *T.parva*-specific, *in vitro* $CD8^+$ T-cell enriched cell-line from animal (A) 592 and (B) 468. Heteroduplex bands representing expanded clonotypes identified from TCRβ chain sequence of individual T-cell clones derived from these cell-line are shown by red arrows, with the number indicating the number of clones out of the ~90 analysed that are represented by each band. Homoduplices are indicated by yellow asterisk.

From each of the two cell-lines approximately 90 $CD8^+$ T-cell clones were generated by limiting dilution and their TCRβ chains sequenced (Chapter 6). Sequence analysis identified 15 TCRβ chains that were expressed by multiple clones (i.e expanded clonotypes) from animal 592 and 10 TCRβ chains that were expressed by multiple clones from animal 468. By comparing the CDR3β-HDA migration patterns of these clones with that of the parent cell-lines it was demonstrated that each of the clonal expansions identified amongst the clones was represented on the CDRβ3-HDA of the bulk line (Figures 3.7 and 3.8).

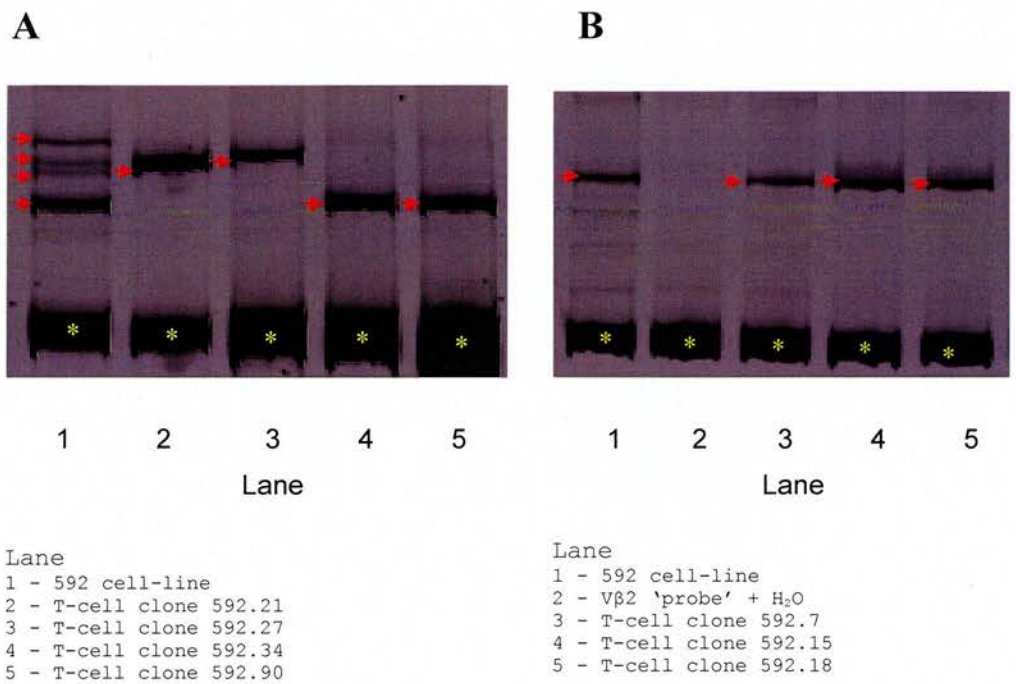


Figure 3.8 Comparison of CDR3β-HDA migration patterns of individual T-cell clones and the parent cell-lines. (A) The heteroduplex bands present in the Vβ6 subfamily CDR3β-HDA of a *T.parva*-specific, *in vitro* $CD8^+$ T-cell enriched cell-line from animal 592 (lane 1a) show co-migration with the heteroduplex bands of the 4 $V\beta6^+$ T-cell clones derived from it (lanes 2a-5a). Clones 592.34 and 592.90 (lanes 4a and 5a) express the same TCRβ chain and co-migrate with the most prominent heteroduplex band in the Vβ6 subfamily CDR3β-HDA of the parent cell-line, suggesting that within Vβ subfamilies CDR3β-HDA is semi-quantitative. Clones 592.21 (lane 2a) and 592.27 (lane 3a) express unique $V\beta6^+$ chains that show co-migration with less prominent heteroduplex band of the 592 cell-line. (B) Examples of the co-migration of the heteroduplex bands present in Vβ2 subfamily CDR3β-HDA of a *T.parva*-specific, *in vitro* $CD8^+$ T-cell enriched cell-line from animal 592. Heteroduplex bands are indicated by red arrows and homoduplices by yellow asterisk.

Many of the other heteroduplex bands present in the cell-lines co-migrated with clones expressing unique TCRβ chains and probably represent additional minor clonal expansions (Figure 3.8a). These results indicate that by using the bovine CDRβ3-HDA it is possible to rapidly acquire an accurate global representation of the clonal composition of complex bovine T-cell populations.

3.4 Discussion

The ability to characterise the clonal composition and TCR repertoires of T-cell responses in humans and mice by molecular analysis of expressed TCRs has been crucial in addressing many important questions relating to T-cell biology. This chapter describes the development and testing of RT-PCR protocols for isolation and identification of expressed bovine TCRβ chains and a CDR3β-HDA technique that can be used to identify clonal expansions within complex bovine T-cell populations. Application of these molecular tools now offers the opportunity to characterise the TCR repertoire and clonal composition of bovine T-cell responses.

Using a sequence database of 46 bovine Vβ genes belonging to 17 subfamilies, 2 sets of essentially Vβ subfamily-specific 5'primers have been designed and validated. Combined use of these two primer sets in an optimised semi-nested PCR protocol to amplify the TCRβ chains expressed by 22 T-cell clones indicated that they were specific for their respective Vβ subfamilies. Although 14 of the 17 Vβ subfamilies were represented amongst these 22 TCRβ PCR products, it must be noted that subfamily specificity has only been established on a limited number of plasmid sub-clones expressing partial TCRβ chains and cDNAs from T-cell clones. It is therefore possible that as yet undetected cross-reactive amplification might occur with (i) members of multiple gene Vβ subfamilies/novel subfamilies not represented within the panel of sub-clones/T-cell clones examined or (ii) full length TCRβ transcripts of the Vβ genes expressed by the sub-clones. The identification of several novel Vβ genes, including a novel subfamily (Vβ5 - see below) indicates that the sequence database used for designing the Vβ subfamily-specific primers does not represent the complete bovine Vβ repertoire. Therefore, the Vβ subfamily-specific snPCR protocol may need re-evaluation and adaptation in the future as additional Vβ genes and subfamilies are identified. The novel Vβ genes identified through the work in this chapter are presented in Chapter 4.

Although the Vβ3$^+$-specific snPCR was shown to generate product from as few as 300 cells, a full titration was not conducted and it is therefore possible that products could be generated from significantly fewer cells. Further evaluation of the sensitivity of the protocol for a variety of subfamilies may be of interest, especially if alternative techniques such as single-cell PCR of tetramer-isolated cells are to be developed in the future (Kedzierska *et al.*, 2004; Maryanski *et al.*, 1999; Turner *et al.*, 2003).

The Vβ subfamily-specific snPCR protocol failed to amplify cDNA from a cloned *T. parva*-transformed cell-line (468TpM) which FACS data had implied was an αβT-cell. This suggested that 468TpM was expressing a novel Vβ gene that lacked annealing sites for the Vβ subfamily-specific primers. Several techniques are available for amplifying TCR transcripts expressing unknown V genes, including inverse PCR using circularised cDNA (Uematsu, 1991), anchored-PCR (Loh *et al.*, 1989) and design of primers that anneal to sequence conserved across V gene subfamilies (Obata *et al.*, 1993). Due to the laborious template modifications that are required in the first two techniques, it was decided to attempt to develop a 'Pan-Vβ' primer. Application of the designed primer in a snPCR protocol successfully amplified the TCRβ chain expressed by 468 TpM, which was identified as member of a novel bovine subfamily (Vβ5).

The sequence of the bovine 'Pan-Vβ' primer designed in this study is identical to that of the human 'Pan-Vβ' primer designed and validated by Obata *et a.l* (1993). This primer has also been used to amplify canine TCRβ chains (Ito *et al.*, 1993), suggesting it may be useful across a spectrum of mammalian species. Furthermore, it has been suggested that the 'Pan-Vβ' primer could be used to amplify TCRα, TCRγ and TCRδ chains (in combination with Cα-, Cγ- and Cδ-specific 3'primers) as the target sequence pentamer (TGGTA) to which the 'Pan-Vβ' primer anneals is also conserved in Vα, Vγ and Vδ genes (Obata *et al.*, 1993). One disadvantage of the 'Pan-Vβ' primer is that it does not generate sequence of the first 100bp of the Vβ gene. However, in humans enough Vβ sequence is obtained to permit identification and differentiation of most Vβ genes (Obata *et al.*, 1993) and from examination of

the available bovine Vβ sequences it appears this also applies to bovine Vβ genes (data not shown).

In subsequent work completed in this study the 'Pan-Vβ' primer enabled identification of another novel Vβ subfamily (VβX - section 4.3.1) and has also been used to amplify the TCRβ chains of a large number of individual T-cell clones, encompassing the majority of the Vβ subfamilies (Chapter 6). However, results from this work (Chapter 6 - data not shown) and that reported by Obata *et al* (1993) revealed one disadvantage of this method, namely that it does not yield readable TCRβ chain sequences from T-cells expressing 2 or more TCRβ chains following direct sequencing of the PCR product. In these circumstances amplification by Vβ subfamily-specific snPCR was required to obtain readable TCRβ chain sequence.

CDR3β-HDA has been used in human and primate studies to provide rapid, global descriptions of the clonal composition of T-cell responses to HIV/SIV (Bernardin *et al.*, 2004; Lopes *et al.*, 2003; Turnbull *et al.*, 2006), EBV (Maini *et al.*, 2000) and BCG vaccination (Bennett *et al.*, 2006). In comparison to the spectratyping technique, CDR3β-HDA has greater sensitivity in the detection of clonally expanded T-cells (Maini *et al.*, 1998; Shen *et al.*, 1998). Furthermore, with spectratyping it is not possible to identify whether disruptions of the Gaussian distribution are due to individual expanded clones or multiple clones expressing CDR3β regions of conserved length (Maryanski *et al.*, 1996). As CDR3β-HDA identifies clones on the basis of CDR3β sequence rather than length, it has a greater capability to differentiate between expanded clonotypes and hence provide greater resolution of the clonal composition of complex αβT-cell populations (although there are limitations to this, as discussed below).

The sensitivity for detection of clones by CDR3β-HDA was assessed for the Vβ2 and Vβ13 subfamilies by seeding *in vitro* derived clones into unfractionated PBMC. Whilst the sensitivity of detection of the $V\beta2^+$ clone was equivalent to that seen in a human CDR3β-HDA technique (Maini *et al.*, 1998) at 1:10,000, the sensitivity of detection of the $V\beta13^+$ clone was reduced by 10-fold to 1:1000 (Figure 3.6). There

are several possible factors that may have contributed to this difference in sensitivity. Subsequent studies described in chapters 4 and 5 demonstrated that the bovine Vβ13 subfamily has undergone substantially more extensive duplication than the Vβ2 subfamily. There may therefore be a higher overall level of expression of Vβ13 in unselected PBMC compared to Vβ2 and consequently more competition from polyclonal $V\beta13^+$ chains during both PCR amplification and heteroduplex formation, resulting in the lower sensitivity in detection of expanded $V\beta13^+$ clones. However, in the absence of comparative quantitative data on expression levels of Vβ13 and Vβ2 (e.g. by quantitative PCR), this remains speculative. The sequence of the individual $V\beta13^+$ and $V\beta2^+$ TCRβ chains expressed by the clones will also influence their ability to form thermodynamically stable heteroduplices with the fluorochrome-labelled 'probes' and so affect sensitivity of detection. Also, as TCR mRNA levels vary widely between different cells (Callan *et al.*, 1998; Maini *et al.*, 1998), it is possible that some of the difference in sensitivity was due to different level of mRNA expression by the seeded $V\beta13^+$ than $V\beta2^+$ clones. Thus, the levels of sensitivity reported in this chapter should not be considered absolute values but rather provide a general indication of sensitivity, which will be subject to both inter- and intra-subfamily variation dependent on a number of variables. It should also be noted that because TCR mRNA levels increase in activated T-cells (Callan *et al.*, 1998; Wotton *et al.*, 1993) it is probable that the sensitivity thresholds identified by use of *in vitro* derived clones in resting PBMC will not be directly applicable to a complex, active T-cell response.

The bovine CDR3β-HDA was used to examine two *T. parva*-specific *in vitro* $CD8^+$ T-cell enriched cell-lines. The results indicated that both cell-lines were complex populations, each containing numerous oligo-clonal T-cell expansions. The clonal composition of these cell-lines as described by CDR3β-HDA was in good agreement with that demonstrated by sequence analysis of the TCRβ chains expressed by approximately 90 clones generated by limiting dilution from each of the cell-lines. All of the expanded clonotypes identified by sequencing of individual clones were represented by a distinct heteroduplex band in the parent cell-lines CDR3β-HDAs (Figure 6.7). Furthermore, most of the clones expressing unique TCRβ chains were

also represent on the parent cell-line CDR3β-HDA (examples shown in Figure 6.8a). Some of the heteroduplex bands did not co-migrate with the TCRβ amplicons from any of the derived clones (e.g. the two Vβ8 heteroduplex bands in the 468 cell-line) and most likely reflect clonal expansions not represented within the individual clones analysed. Direct sequencing of the bands from the gel (Bernardin *et al.*, 2003, this study - data not shown) could be used to confirm that they are genuine heteroduplices. Alternatively these bands may be artefactual. Use of proof-reading Taq polymerase for PCR amplification in a human CDR3β-HDA protocol helped in reducing the presence of artefactual bands (Dr. A. Bennett, personal communication) and may be something to consider using in future work with the bovine CDR3β-HDA.

In general, the intensity of heteroduplex bands evident within subfamilies appears to reflect the size of the clonal expansions represented (e.g. Figure 6.8a). However, the intensity of heteroduplex bands is an unreliable parameter for inter-subfamily comparison of clone sizes. For example, although the larger Vβ28$^+$ clonotype is represented by the most intense heteroduplex band on the 592 cell-line CDR3β-HDA (upper heteroduplex band in final lane of Figure 3.7a), it is notable that the band representing the Vβ17b$^+$ clonotype (comprising 6 individual clones - Figure 3.7a) is fainter than that representing the Vβ8$^+$ clonotype (comprising only 2 individual clones - Figure 3.7a). This is unsurprising given the inter-subfamily variation in sensitivity that, as discussed above, is inherent in the technique - but it does have implications for the accurate interpretation of the results of CDR3β-HDA.

The ability of CDR3β-HDA to separate and allow identification of expanded clonotypes is a function of the expression of unique CDR3β sequences by individual clonotypes. However, it is well documented that some clonotypes sharing the same pMHC-specificity can express TCRs that have TCRβ chains that are either identical or show minimal differences (Annels *et al.*, 2000; Hamrouni *et al.*, 2003; Kedzierska *et al.*, 2004) and that subsequently may not be effectively segregated in CDR3β-HDA (Cook and Beverley, 2001). Amongst the 592 clones there were two Vβ28s1$^+$ clonotypes that could be differentiated on the basis of (i) expression of different

alleles of the Vβ28s1 gene and (ii) co-expression in one clonotype of a non-functional Vβ13s6^{+} TCRβ chain (section 6.3.6.2), but which expressed identical CDR3β sequences and consequently shared the same CDR3β-HDA migration pattern. Therefore, in the uncloned 592 cell-line, both clonotypes were represented by the same heteroduplex band (Figure 3.7 last lane, upper band). The ability to segregate these two clonotypes by sequencing the TCRβ chains of individual clones was to some extent fortuitous and is a reminder that care is needed in evaluation of clonal composition of complex T-cell populations by CDR3β-HDA (and other techniques that employ analysis of TCRβ sequence as a marker of clonality).

In summary, a suite of molecular tools that allow the sequencing of bovine TCRβ chains and also the accurate and rapid characterisation of the clonal composition of complex bovine T-cell populations has been developed and validated. These generic tools can now be applied to study the TCR repertoire and clonal composition of a wide range of bovine T-cell responses, including those induced by *T. parva*.

4 Chapter 4 – Examination of the bovine TCRβ gene repertoire

4.1 Introduction

The αβTCR repertoire is characterised by enormous diversity (Arstila *et al.*, 1999; Casrouge *et al.*, 2000) which facilitates the recognition of the myriad pMHC ligands that may be presented to T-cells. Much of this diversity is generated by somatic recombination, whereby different permutations of individual non-contiguous variable (V), diversity (D - in β chain only) and joining (J) gene segments, selected essentially at random from the multiple copies available, are brought together to form the single exons that encode the variable domains of the α and β chains. Further diversity is created by (i) nucleotide editing at the V(D)J junctions, with terminal nucleotides removed by exonuclease activity and non-germline nucleotides added by terminal deoxynucleotide transferase (Tdt) and (ii) different combinations of α and β chains.

Following extensive analysis of TCRβ chain transcripts during the 1980 and 1990s (Arden *et al.*, 1995a, b) and genomic sequencing of the TCRB loci (Rowen *et al.*, 1996, GenBank accession numbers AE000663-AE000665), the full repertoire of human and murine TCRβ genes is known. In other species such as cattle (Houston, 1997; Houston and Morrison, 1999; Tanaka *et al.*, 1990), pig (Baron *et al.*, 2001; Butler *et al.*, 2005), horse (Schrenzel *et al.*, 1994), goat (Obexer-Ruff *et al.*, 1998), sheep (Halsey *et al.*, 1999), rabbit (Isono *et al.*, 1994) and rat (Smith *et al.*, 1991) partial information on TCRβ genes, of varying levels of resolution has also been gathered from cDNA analyses.

TCRβV (Vβ) genes are classified into subfamilies, the members of which share >75% nucleotide sequence identity (Kazatchkine, 1995). Inter-species comparisons reveal a much higher level of sequence similarity between Vβ gene in orthologous subfamilies in different species than between Vβ genes of different subfamilies in the same species (Baron *et al.*, 2001; Clark *et al.*, 1995) reflecting establishment of some Vβ subfamilies prior to mammalian radiation. However, both the number of Vβ

subfamilies and the number of members within subfamilies differ between humans and mice (section 1.2.3.2), indicating that the ancestral Vβ genes have been subject to different deletion and duplication events in the 80-110 million years subsequent to human-mouse divergence (Kumar and Hedges, 1998; Li *et al.*, 1990). Examination of the available data for other species also demonstrates that during speciation the repertoires of individual species have been moulded by different duplication events. For example, Vβ22 and Vβ1, both single-member subfamilies in humans and mice, have multiple members in pig and sheep respectively (Baron *et al.*, 2001; Halsey *et al.*, 1999). As with the Vβ genes there is also a high level of inter-species sequence conservation of the Dβ, Jβ and Cβ genes (Baron *et al.*, 2001; Schrenzel *et al.*, 1994; Williams *et al.*, 1991), implying that the basic Dβ, Jβ and Cβ repertoires were also formed early in mammalian speciation.

Characterisation of the full complement of the TCRβ genes in human and mice has allowed the development of methods to comprehensively study TCRβ expression in these species, yielding useful information on various aspects of the clonality and TCR repertoire of antigen-specific T-cells (section 3.1) Application of such methods to farm animal species is impeded by incomplete knowledge of the TCRβ gene repertoires. Previous studies of cattle TCRβ genes have identified 46 Vβ genes distributed within 17 subfamilies and 13 Jβ genes (Buitkamp *et al.*, 1993, Russell and MacHugh, unpublished; Houston, 1997; Houston and Morrison, 1999; Tanaka *et al.*, 1990). Two of the subfamilies, Vβ1 and Vβ2 have been found to be extensively duplicated, having 11 and 7 members respectively (Houston, 1997).

In the course of developing techniques to analyse bovine TCRβ chain sequence (Chapter 3) and their application to study the $CD8^+$ T-cell response to *Theileria parva* (Chapter 6), a large dataset of rearranged TCRβ chain cDNA sequences has been generated. Analysis of this dataset identified 22 novel Vβ genes, including members of 3 novel subfamilies, and 4 novel Jβ genes. Notably, Vβ4, 10 and 13 were identified as additional subfamilies that had undergone extensive duplication and subsequent subfamily-specific cDNA analysis identified further novel genes within these subfamilies. The results presented in this chapter contribute to a more

thorough characterisation of the bovine TCRβ gene repertoire which will be useful in developing and refining molecular techniques to study bovine T-cell responses.

any of the previously identified Vβ genes were, according to convention (section 2.5.1), considered to represent a novel gene and added to the database.

Using the IMGT/V-QUEST software programme (Guidicelli et al 2004 - http://imgt.cines.fr) the human Vβ gene with the highest nucleotide similarity to each Vβ gene was identified and the subfamily of the bovine Vβ genes thus determined (except for bovine VβX - section 4.3.1). The IMGT/V-QUEST programme was also used to align the bovine Vβ genes according to the IMGT unique numbering system and define the complementarity determining regions (CDRs).

Sequence data for porcine VβX and murine Vβ2, were obtained from sequences AY690918 and AE000663 deposited in the NCBI nucleotide database (http://www.ncbi.nlm.nih.gov) - Appendix B.

(ii) Analysis and nomenclature of Jβ genes

Only sequence 5' to the Cβ gene up to and including that encoding the conserved Phe-Gly-X-Gly (F-G-X-G) motif of Jβ genes has been considered to represent germline Jβ sequence (sequence 5' to this could have been altered during somatic recombination). TCRβ transcripts were accordingly edited prior to further analysis.

Edited Jβ gene sequences were compared with 'in-house' databases containing the human and known bovine Jβ gene nucleotide sequences using the BLASTN algorithm. Human Jβ sequences were derived from annotated genomic sequence of the human TCRB locus (GenBank accession number U66061), and bovine sequences from Tanaka *et al* (1990) and Houston (1997). In the WHO-IUIS system Jβ gene nomenclature is determined by genomic location (Kazatchkine, 1995). Therefore, the bovine Jβ genes were assigned temporary numerical names prior to the annotation of the Jβ genes in the bovine TCRB locus (section 5.3.2).

Percentage nucleotide/amino acid identities were calculated using the Lasergene software (DNAstar, Madison, WI)

4.2.3 Subfamily-specific generation and sequencing of TCRβ chain transcripts

4.2.3.1 Subfamily-specific PCR amplification of cDNA

Naïve PBMC from animals 641 and 605 were isolated (section 2.3.1) and cDNA generated as described in sections 2.4.1-2.4.3. In separate reactions cDNA was amplified with the VB4ext, VB7ext, VB10ext and VB13ext 5'primers (Table 2.2) in combination with the common Cβ-specific 3'primer BCext (Table 2.3). Each PCR reaction used 20 pmol of each primer, 1 unit BIOTAQ (5units/µl Bioline, London, UK), 4 µl SM-0005 buffer (ABgene, Epsom, Surrey, UK – Appendix C.7), 2 µl cDNA, and nuclease free water to give a final volume of 40 µl. The programme used was as follows – 5 min at 94°C, 5 cycles of (1 min at 94°C, 1 min at 60°C, 1 min at 72°C), 25 cycles of (30 s at 94°C, 1 min at 60°C, 1 min at 72°C) and a final extension period of 5 min at 72°C.

4.2.3.2 Cloning and sequencing of TCRβ chain transcripts

Following confirmation of PCR products of the expected size by agarose gel electrophoresis (section 2.4.5), the products were purified using the Wizard PCR Preps DNA Purification System (section 2.4.6), sub-cloned into plasmid pGEM-T Easy vectors and used to transform High Efficiency Competent JM109 cells (section 2.4.7). The presence of inserts of the anticipated size in selected sub-clones was verified by PCR as described in section 2.4.7.1. The resulting PCR products were sequenced (section 2.4.9.1-2.4.9.2) and the sequences analysed as described in section 4.2.2.

4.2.4 Southern blot

4.2.4.1 Culture of bovine fibroblasts

A bovine skin fibroblast line from animal 605 was cultured in supplemented D-MEM medium (Appendix C.20) incubated at 37°C in a humidified atmosphere of 5% CO_2 in air. Cells grown in culture flasks were washed once in PBS (Appendix C.6),

incubated at 37°C for 5 - 10 min in trypsin (Trypsin EDTA (1x), Gibco, Paisley, UK) and the harvested cells washed in foetal bovine serum (Gibco, Paisley, UK), counted and then stored at -20°C until use.

4.2.4.2 Extraction and quantification of genomic DNA

Genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA) according to the manufacturer's instructions. Briefly, 5 x 10^7 pelleted cells were washed in PBS, re-suspended and then lysed in 3 ml of Nuclei Lysis Solution. 30 µl of RNAse Solution was mixed into the nuclear lysate and the mixture incubated at 37°C for 15 – 30 min. 2 ml of Protein Precipitation Solution was added and the mixture incubated on ice for 5 min prior to centrifugation at 13,000-16,000 *x g* for 4 min. The supernatant (containing the DNA) was carefully transferred to a clean 15 ml tube, 6 ml of isopropanol added at room temperature and the tube inverted gently until white threads of DNA became visible. The DNA was pelleted by centrifugation at 15,000 *x g* for 1 min at room temperature, washed in 6 ml of 70% ethanol, air-dried and re-hydrated by incubation in 500 µl of nuclease-free water at 65°C for 1 h and subsequently overnight at room temperature. The re-suspended DNA was quantified as described in section 2.4.2 and then stored at 2 - 8°C until use.

4.2.4.3 Enzyme digestion and separation of DNA fragments by agarose gel electrophoresis

Aliquots of 20 µg of genomic DNA were incubated overnight at 37°C in 200 µl of the respective enzyme digestion reactions containing *Hind*III, *Xba*I (10U/µg DNA), *Sph*I or *Ssp*I (2.5U/µg DNA) in the relevant buffer and, where required, bovine serum albumin (BSA). All enzymes, buffers and BSA were from New England Biolabs (Hitchin, Herts. UK). To enable subsequent loading into an agarose gel, the digested DNA samples were concentrated by ethanol precipitation; 400 µl ethanol and 20 µl 3M sodium acetate pH 5.6 (Appendix C.19) was added to each digest reaction. Following incubation at -20°C for 30 min, digested DNA was pelleted by centrifuging at 12,000 *x g* for 15 min at 0°C. Digested DNA pellets were washed

twice in 70% ethanol, air dried and re-suspended overnight at room-temperature in 25 µl of nuclease-free water.

Digested DNA was loaded into a 20cm x 20cm 0.8 % agarose gel (GNA 200 system, Pharmacia Biotech AB, Uppsala, Sweden) prepared with 1 x TAE (Appendix C.9) and run at 33 V for 16 h using 1 x TAE as running buffer. 5 µg of Lambda DNA digested with *Hind*III (Promega, Madison, WI, USA) was loaded into each gel to provide molecular weight markers. After electrophoresis, the gels were stained in ethidium bromide and visualised as described in section 2.4.5.

4.2.4.4 Transfer of DNA from agarose gels to nylon membranes

DNA was transferred from the agarose gel to a nylon membrane (Hybond-N+, Amersham Biosciences, Chalfont St. Giles, UK) by capillary transfer as described in (Sambrook *et al.*, 1989). To facilitate the transfer, the DNA first had to be depurinated and denatured. DNA was depurinated by agitating the gel for approximately 30 min whilst submerged in 0.125M HCl (Appendix C.21), the gel was washed in distilled water and then the DNA denatured by agitating the gel for approximately 45 min whilst submerged in 1.5M NaCl, 0.5M NaOH (Appendix C.22). After a further wash, the gel was neutralised by agitating for approximately 30 min whilst submerged in 1M Tris (pH7.4) containing 1.5M NaCl (Appendix C.23). Following a further wash, the gel was placed in a capillary blot for 16 h, with 20 x SSC (Appendix C.24) used as the transfer buffer. Following transfer, the DNA was fixed to the membrane by incubating for 2 h at 80°C whilst enclosed in dry 3MM Whatman paper.

4.2.4.5 Probe generation

PCR amplification of the Vβ regions of plasmid clones VB1.1, VB2.2, VB13.1 and VB14.6 (Table 3.3) provided DNA to generate Vβ1, 2, 13 and 14 subfamily-specific probes. Individual 100 µl reactions contained 50 pmol of each of the relevant 5' and 3'primers (Table 4.1), 2.5 units BIOTAQ (5units/µl Bioline, London, UK), 10 µl SM-0005 buffer (ABgene, Epsom, Surrey, UK – Appendix C.7), 5 µl plasmid DNA

(10ng/ml) and nuclease free water to give a final volume of 100 µl. Reactions were completed using the following programme – 3 min at 94°C, 30 cycles of (94°C for 1 min, 54°C for 1 min and 72°C for 1 min) and a final extension period of 3 min at 72°C. Analysis of PCR products by electrophoresis (section 2.4.5) confirmed generation of products of the anticipated size. PCR products were purified using Wizard PCR Preps DNA Purification System (section 2.4.6) and quantified (section 2.4.2).

Primer	Sequence
VB1*1for	CCC CAA ATA (CT)(CT)T GAT CAA ATC AAG
VB1*1rev	GCT GGC ACA GAG ATA CAC G
VB2*1for	CTC TCG TCT CTC AGC AGC
VB2*2rev	GCT GTC TGC AGG ATC CAC
VB13*1for	CAC TCA GGA CCC CAG ATT
VB13*1rev	AGT ACA CAG ATG TCT GGG AG
VB14*2for	AGG CCC CCT GGA (AT)GC CGA T
VB14*1rev	GGC CAA CTC CAA GGT CAG

Table 4.1 Forward and reverse primers used for generation of Vβ1, 2, 13 and 14 subfamily specific DNA probes.

Each DNA was labelled with ^{32}P-dCTP using the Random Primed DNA Labelling Kit (Roche Applied Science, Penzberg, Germany) according to the manufacturer's instructions. Briefly, nuclease-free water was added to 25 ng aliquots of each PCR product to give a final volume of 9 µl. The DNA was denatured by heating for 10 min at 95°C (and then immediately placed on ice to prevent re-naturing) and the following added – 1 µl each of dATP, dGTP, dTTP (all at 0.5 mM), 2 µl Reaction Mixture (hexanucleotide mixture in 10x reaction buffer), 2 units Klenow Enzyme (2 U/µl) and 50 µCi (α-^{32}P) dCTP 3000Ci/mmol (GE Healthcare Ltd., Little Chalfont, Bucks, UK). The reaction was mixed and briefly centrifuged before incubation at 37°C for 30 min. The reaction was stopped by the addition of 2 µl 0.2M EDTA (pH8.0) and heating to 65°C for 10 min. Immediately before hybridization probes were denatured by heating at 95 - 100°C for 2 - 5 min and incubated on ice until used.

4.2.4.6 Hybridization and visualisation

Blots were pre-hybridised by complete submersion in pre-warmed Rapid-Hyb buffer (Roche, Applied Science, Penzberg, Germany) for 15 min at 65°C before the probes were added to give a final concentration of approximately 2 ng of probe DNA per ml of Rapid-Hyb buffer. Hybridisation was completed at 65°C for 2 h with continuous rotation. After hybridisation, blots were given three stringency washes; the first in 2 x SSC, 0.1%(w/v) SDS for 20 min at room temperature, the second in 0.5 x SSC, 0.1%(w/v) SDS for 15 min at 65°C and the third in 0.1 x SSC, 0.1%(w/v) SDS for 15 min at 65 °C. The blots were moistened in 0.1 x SSC, 0.1%(w/v) SDS, wrapped in Saran wrap and exposed to x-ray film (Biomax MS Film, Kodak, Hemel Hempstead, Herts, UK) whilst incubated at -70 °C. Exposures for different times ranging from 1 hour to 16 hours were made.

4.3 Results

4.3.1 Analysis of rearranged bovine TCRβ chains identifies 22 novel Vβ genes and 3 new Vβ subfamilies.

Within the dataset of approximately 450 bovine TCRβ chain sequences generated during the completion of the work described in Chapters 3 and 6, 173 unique TCRβ chain rearrangements were identified. The currently known repertoire of 46 bovine Vβ genes has been established from the analysis of a combined total of ~150 TCRβ chain sequences in several studies (Buitkamp *et al.*, 1993, Russell and MacHugh, unpublished; Houston, 1997; Houston and Morrison, 1999; Tanaka *et al.*, 1990). The availability of 173 new bovine TCRβ chain sequence thus offered a significant opportunity to re-evaluate the repertoire of bovine Vβ genes.

Amongst the 173 rearranged TCRβ transcripts, 57 Vβ genes distributed within 20 subfamilies were identified. This included 35 of the 46 Vβ genes identified in previous studies including 6 that had been represented by a single sequence result (Vβ1s11, 4s3, 7s2, 8s2, 9s1, 16s1). All previously identified subfamilies were represented within the dataset; the 11 previously identified genes not represented were all members of multi-member subfamilies (Vβ1s2, 1s4, 1s10, 2s1, 2s2, 2s5, 4s1, 4s2, 4s4, 10s1 and 10s2).

The other 22 Vβ sequences identified represented novel Vβ genes. Eighteen of these genes were new members of already established Vβ subfamilies, including 2 new members each for subfamilies Vβ1 (1s12, 1s13 - Figure 4.1), Vβ2 (2s8, 2s9 - Figure 4.2) and Vβ4 (4s5, 4s6 - Figure 4.3) and a single new Vβ17 gene (17s3 - Figure 4.4). More notably, numerous novel genes were found in subfamilies Vβ10 (10s4, 10s5, 10s6, 10s7 - Figure 4.5) and Vβ13 (13s4, 13s5, 13s6, 13s7, 13s8, 13s9, 13s10 - Figure 4.6), which respectively more than double and triple the number of identified members of these 2 subfamilies and expands them to sizes comparable with subfamilies Vβ1 and Vβ2. As would be anticipated, each of these novel subfamily members had ≥75% nucleotide sequence identity with other members of their respective subfamilies (Figures 4.1c- 4.6c). From the partial cDNA sequences

obtained 2 of the Vβ genes (Vβ10s6 and Vβ13s6) would be predicted to encode non-functional Vβ polypeptides. Vβ10s6 is predicted to be a pseudo-gene due to a frameshift in the reading frame and Vβ13s6 is predicted to be an open reading frame (ORF) pseudo-gene due to loss of the codon for the conserved cysteine residue at position 104 which is essential for the correct protein folding of the Vβ immunoglobulin domain.

The four remaining genes were members of novel bovine Vβ subfamilies. The genes designated as bovine Vβ5s1, Vβ12s1 and Vβ12s2 were classified as such because they showed highest nucleotide sequence identity to human Vβ5s3 (73.6%) and human Vβ12s1 (74.6% for both Vβ12s1 and 12s2) respectively (sections 1.2.2 and 2.5.1). However, both Vβ12 genes have nearly as high nucleotide identity with members of the human Vβ13 subfamily (Vβ12s1 has 74.2% nucleotide identity with human Vβ13s1 and Vβ12s1 has 73.8% nucleotide identity with human Vβ13s5), and show sufficiently high homology (>75%) to all bovine Vβ13 subfamily members, that if considered without reference to human sequence comparison, both Vβ12 genes would be incorporated within the Vβ13 subfamily (Figure 4.6c). Furthermore, both Vβ12 genes were identified following amplification with Vβ13 subfamily-specific primers. Similarly, Vβ5s1 shows higher nucleotide identity to human Vβ1 (70.4%) than some human Vβ5 subfamily members (e.g. 67.4% identity with human Vβ5s1) and displays >75% nucleotide homology to 7 out of the 13 members of the bovine Vβ1 subfamily (Figure 4.1c). The ambiguity in defining bovine genes as members of Vβ1 or 5 and Vβ12 or 13 reflects the situation in humans where members of these subfamilies frequently share >75% homology (Arden *et al.*, 1995a). The nucleotide and amino acid sequences of Vβ5s1 are shown aligned with the novel Vβ1 subfamily genes in Figure 4.1 and those of Vβ12s1 and 12s2 are shown in Figure 4.6 with the novel Vβ13 sequences.

The final novel gene identified had no human orthologue (<30% nucleotide identity with any human Vβ gene) but shared 76.2% nucleotide identity with murine Vβ2. A *Sus scrofa* (porcine) Vβ gene identified by Baron *et al* (2001) in a miniature pig showed similar phylogenetic characteristics (GenBank accession number AY690918)

and was subsequently given the subfamily designation VβX by Butler *et al.* (2005). The bovine gene has 80.3% nucleotide identity with the porcine VβX gene, and was given the corresponding name - VβXs1. The nucleotide and amino acid sequences of bovine VβXs1 are given in Figure 4.7.

4.3.2 Further examination of large multi-member subfamilies

Nearly 40% of the Vβ genes identified in the above analysis (22/57) were novel, with all but 2 (Vβ17s3 and VβXs1) belonging to one of five large subfamilies - Vβ1, Vβ2, Vβ4, Vβ10 and Vβ13 or the 'associated' Vβ5 and Vβ12 subfamilies. The high frequency with which novel genes were identified in the large subfamilies, especially in Vβ10 and Vβ13, suggested that the complement of known genes belonging to these subfamilies was still incomplete.

The ability to amplify TCRβ chains by PCR using Vβ subfamily-specific primers offered the opportunity to analyze the membership of selected subfamilies. In separate reactions cDNA from unfractionated PBMC was amplified using the VB10ext and VB13ext 5'primers in combination with the Cβ-specific 3'primer BCext and the products sub-cloned into plasmid vectors. Analysis of the resulting partial TCRβ chain transcripts revealed 74 unique Vβ10 TCRβ chain rearrangements (derived from animal 641) and 161 unique Vβ13 TCRβ chain rearrangements (77 derived from animal 605 and 84 from animal 641). These sequences included four novel Vβ10 genes (Vβ10s8, 10s9, 10s10 and 10s11 - Figure 4.5) and nine novel Vβ13 genes (Vβ13s11, 13s12, 13s13, 13s14, 13s15, 13s16, 13s17, 13s18, 13s19 - Figure 4.6). By including the previously identified sequences, this takes the total membership of subfamily Vβ10 to 11 genes and Vβ13 to 19 genes. However, (i) not all of the previously identified Vβ10 and Vβ13 genes were present in this set of sequences (Vβ10s5, 10s6, 13s3, 13s7 and 13s9 were absent), (ii) there were marked differences in the representation of different Vβ13 members in the sub-clones derived from animals 641 and 605 (Figure 4.8), with some members only present in the sub-clones from one animal (e.g. Vβ13s16 in animal 641), and (iii) several genes

Figure 4.1 A) Nucleotide sequence and B) amino acid sequence of novel Vβ1 and Vβ5 genes. Sequences are aligned and CDR1 and CDR2 (shaded regions) defined according to the IMGT unique numbering system. Identity is shown as dashes and gaps by dots. C) Percentage nucleotide (top-right) and amino acid (bottom-left) sequence identity of the Vβ1 and Vβ5 genes. Shaded boxes indicate >75% nucleotide identity between Vβ5s1 and Vβ1 subfamily members - the value conventionally used to define members of the same subfamily.

A

```
        1           11          21          31          41          51
        .           .           .           .           .           .
VB1s12                                                                GGACC
VB1s13
VB5s1

                                                CDR1
       61          71          81          91         101         111
        .           .           .           .           .           .
VB1s12 CTGGGATGCT CCCCTGTCTC CGGACACCTC TCC....... .......... ....GTGTCC
VB1s13
VB5s1

                                                                    CDR2
      121         131         141         151         161         171
        .           .           .           .           .           .
VB1s12 TGGTACCAAC AGGCCCTGGG CCAGAGCCCC GAGTTCCTCA CTCAGTATTA CAGGCAGGAA
VB1s13      ----- ---------- ----G----- C-------T- T---------- T-ATGG-A--
VB5s1       ----- ------A-A- T---G----- C-----A--T T-G---T--- TGAAACTCTG

      181         191         201         211         221         231
        .           .           .           .           .           .
VB1s12 GTT....... .....AGGGG AGAAGCACAG CTCCCG...G ATCGATTCTC AGGAAAACAG
VB1s13 GAG....... .....-AT-A GA---G-A-C A-G--A...- ---------- ---TG-----
VB5s1  CAA....... .....--A-C TA---G-A-C T-TT--...A ----G----T --C-------

      241         251         261         271         281         291
        .           .           .           .           .           .
VB1s12 TTT...AGTG ACTTTCACTC TGAGCTGAAC TTGAGCTCCT TGGAGCTGAC GGACTCAGCC
VB1s13 --C...---- ---C-AG--- -C--A----- C--------- ---------- A--------T
VB5s1  --C...CCT- -----AG--- ---------- G---A----- ----C--A-- ------T---


      301         311
        .           .
VB1s12 GTGTATCTCT GT
VB1s13 ---------- --
VB5s1  C----C---- --
```

B

```
                                        CDR1                            CDR2
       1           11          21          31          41          51
       .           .           .           .           .           .
VB1s12 ---------- ---------T LGCSPVSGHL S.......VS WYQQALGQSP EFLTQYYRQE
VB1s13                                                -------G -Q-I---NGK
VB5s1                                                 ---QS-G- Q-IFEF-ETL

       61          71          81          91          101
       .           .           .           .           .
VB1s12 V....RGEAQ LP.DRFSGKQ F.SDFHSELN LSSLELTDSA VYLC
VB1s13 E....NEKGN M-.-----E- -.--SS-QM- ---------- ----
VB5s1  Q....-AKGN FS.N--LA-- -.P--S---- VN--D----- L---
```

C

Nucleotide sequence identity (%)

Amino acid sequence identity (%)

	1s1	1s2	1s3	1s4	1s5	1s6	1s7	1s8	1s9	1s10	1s11	1s12	1s13	5s1
1s1		95.6	80.2	81.7	87.6	80.2	83.5	84.5	80.3	81.0	80.6	93.1	80.5	73.8
1s2	90.1		80.6	80.6	88.4	79.1	84.6	83.8	80.0	80.2	80.1	96.3	80.6	72.6
1s3	65.9	68.1		85.3	84.2	87.2	90.5	85.3	88.2	89.2	93.0	80.3	87.0	74.8
1s4	70.3	70.3	79.1		85.0	92.6	89.4	86.3	88.5	86.0	86.9	79.8	92.9	74.3
1s5	78.7	80.1	75.2	77.3		84.2	86.9	87.3	83.9	82.3	85.7	87.2	83.4	75.7
1s6	67.0	67.0	81.3	89.0	74.2		93.8	82.8	86.7	87.9	89.2	78.9	97.0	75.2
1s7	72.5	78.0	86.8	83.5	79.8	89.0		87.5	88.9	88.3	91.2	85.9	93.5	76.7
1s8	76.7	74.4	74.4	78.9	78.7	72.2	80.0		84.5	82.7	86.1	82.6	83.4	74.3
1s9	70.0	71.1	84.4	80.0	75.2	74.4	82.2	77.8		84.8	86.5	78.4	87.0	75.1
1s10	64.9	66.2	83.1	77.9	72.7	80.5	80.5	70.1	76.6		94.8	80.7	87.6	75.7
1s11	64.0	67.4	88.4	79.1	75.6	82.6	87.2	73.2	79.0	87.5		81.7	88.2	76.3
1s12	87.5	95.8	65.3	70.4	79.2	65.3	77.5	73.6	69.4	68.0	66.7		79.8	72.7
1s13	64.2	69.6	80.3	87.5	73.2	94.6	91.1	73.2	71.4	80.3	82.1	67.8		75.3
5s1	53.6	53.5	55.4	51.7	53.6	53.6	58.9	55.4	53.4	58.9	58.9	51.8	53.6	

Figure 4.2 A) Nucleotide sequence and B) amino acid sequence of novel Vβ2 genes. Sequences are aligned and CDR1 and CDR2 (shaded regions) defined according to the IMGT unique numbering system. Identity is shown as dashes and gaps by dots. C) Percentage nucleotide (top-right) and amino acid (bottom-left) sequence identity of the Vβ2 genes.

A

```
      1           11          21          31          41          51
      .           .           .           .           .           .
VB2s8                         CCCACAGG GCTTTCAGTA AGAGTGGAGC CTCTGTGACC
VB2s9                         ---G---- ---------- -------G-- ----------

                                                CDR1
     61          71          81          91         101         111
      .           .           .           .           .           .
VB2s8 ATCGAGTGCC GTGCAGTGGA CTTTCAAGCC TCAAGT.... .......... ....ATGTTC
VB2s9 ---------- -----C---- ---------T A---C-.... .......... ....T-----

                                                                  CDR2
      121         131         141         151         161         171
       .           .           .           .           .           .
VB2s8  TGGTACCGTC AGTTCCCAAA ACGAGGCCTC GTGCTGATGG CAACTTCTAA TGTGGGCACT
VB2s9  ---------- -------G-- ---------- ---------- ---------- --C-------

      181         191         201         211         221         231
       .           .           .           .           .           .
VB2s8  AATGCC.... .....ACATA CGAACAAGGT TATACCAAGG ACAAGTTTCC CATCAGTCAT
VB2s9  ------.... .....----- ---------- ----A----- ---C------ ----------

      241         251         261         271         281         291
       .           .           .           .           .           .
VB2s8  CCA...GACC TAACATTTTC ATCTCTGATG GTGACAAATG TGGATCCTGC AGAGAGCGGC
VB2s9  ---...---- ---------- ---------- ---------- ---------- ----------

      301         311
       .           .
VB2s8  CTCTACTTCT GT
VB2s9  ---------- --
```

B

```
                                        CDR1                                CDR2
      1          11          21          31          41          51
      .          .           .           .           .           .
VB2s8         HR AFSKSGASVT IECRAVDFQA SS......MF WYRQFPKRGL VLMATSNVGT
VB2s9         R- ---------- -----L---- TT......V- ---------- -------A--

      61         71          81          91          101
      .          .           .           .           .
VB2s8 NA...TYEQ GYTKDKFPIS HP.DLTFSSLM VTNVDPADSG LYFC
VB2s9 --...---- --N--T---- --.-------- ---------- ----
```

C

Nucleotide sequence identity (%)

Amino aicd sequence identity (%)

	2s1	2s2	2s3	2s4	2s5	2s6	2s7	2s8	2s9
2s1		95.3	96.4	95.3	96.4	94.3	96.5	91.9	93.8
2s2	89.4		97.1	94.6	96.8	91.2	96.1	96.9	95.0
2s3	93.6	91.5		93.8	96.8	96.4	95.3	93.8	92.7
2s4	92.4	90.0	88.0		96.0	96.7	95.6	93.5	95.4
2s5	92.5	90.4	90.4	94.6		97.2	96.8	94.2	95.4
2s6	87.2	91.5	89.4	95.6	91.4		95.4	95.4	94.6
2s7	94.5	92.5	90.4	95.6	94.7	90.4		93.8	96.5
2s8	82.6	94.2	84.9	86.0	86.0	88.4	87.0		95.8
2s9	87.2	89.5	82.6	91.8	90.7	87.2	91.9	90.7	

Figure 4.3 A) Nucleotide sequence and B) amino acid sequence of novel Vβ4 genes. Sequences are aligned and CDR1 and CDR2 (shaded regions) defined according to the IMGT unique numbering system. Identity is shown as dashes and gaps by dots. C) Percentage nucleotide (top-right) and amino acid (bottom-left) sequence identity of the Vβ4 genes.

A

```
      1          11         21         31         41         51
      .          .          .          .          .          .
VB4s5                                  CAGG GCCATCTGTC AACGCGGGAC CTCCGTGATG
VB4s6                                  ---- ---------- ---------- ----A---C-
VB4s7                                  ---- ---------- -----A---- ----------
VB4s8                                  ---- ---------- ---------- ----------

                                            CDR1
     61         71         81         91        101        111
      .          .          .          .          .          .
VB4s5 ATCGAGTGTC GTCAGGTTGA TAGCCAGCTC ACCGCG.... .......... ....ATGTAC
VB4s6 ---A------ ...-A--C-- ---T------ ---TG-.... .......... ....------
VB4s7 ---------- GTC------- ---------- ---AT-.... .......... ....------
VB4s8 ---------- ...----C-- --------G- --TTT-.... .......... ....------

                                                                  CDR2
     121        131        141        151        161        171
      .          .          .          .          .          .
VB4s5 TGGTACCGTC AGCTTCCAGG ACAGAGCTTG GTGCTGATGG CTACTGCCAA TCAGGGCTCC
VB4s6 --------A- ---------- ---------- ---------- ---------- ----------
VB4s7 ---------- ---------- ---------- ---------- ---------- ----------
VB4s8 ---------- ---------- ---------- ---------- ---------- ----------

     181        191        201        211        221        231
      .          .          .          .          .          .
VB4s5 AAGGCT.... .....ACTTA CGAGAGTGGG TTTACTGACA ACAAATTTCC CATTAGCCGC
VB4s6 ------.... .....----- T--------- --------GG ----G----- ----GA----
VB4s7 ------.... .....----- ---------- -----C--GG ----G----- ----GA----
VB4s8 ------.... .....----- ---------- --------GG ----G----- ----GA----

     241        251        261        271        281        291
      .          .          .          .          .          .
VB4s5 CCG...GTCC TTGCGTTCTC AACTCTGACT GTGAGCAACG CGAGCTCCGA AGACAGCAGC
VB4s6 ---...AAA- -G-A------ ---C------ ---------- ---------- ------T---
VB4s7 ---...AAA- -G-A------ ---------- --------T- ---------- ---T--T---
VB4s8 ---...AAA- -G-A-A---- ---------- ---------A ---------- ------T---
```

```
      301         311
       .           .
VB4s5 TCTTATTTCT GC
VB4s6 ---------- --
VB4s7 ---------- --
VB4s8 ---------- --
```

B

```
                                        CDR1                          CDR2
      1          11          21          31          41          51
      .           .           .           .           .           .
VB4s5          R AICQRGTSVM IECRQVDSQL TA......MY WYRQLPGQSL VLMATANQGS
VB4s6          - --------MT -K-.------ -W......-- ---------- ----------
VB4s7          - -----R---- ---------- -M......-- ---------- ----------
VB4s8          - ---------- ---.-----V NF......-- ---------- ----------

      61          71          81          91          101
      .           .           .           .           .
VB4s5 KA...TYESG FTDNKFPIS RP.VLAFSTLT VSNASSEDSS SYFC
VB4s6 --...----- --ED----D --.K-E----- ---------- ----
VB4s7 --...----- --ED----D --.K-E----- ---------- ----
VB4s8 --...----- --ED----D --.K-EI---- ---T------ ----
```

C

Nucleotide sequence identity (%)

Amino acid sequence identity (%)

	4s1	4s2	4s3	4s4	4s5	4s6	4s7	4s8
4s1		90.3	96.4	90.8	92.9	90.1	91.0	95.2
4s2	83.5		93.2	89.2	88.7	95.3	92.5	93.7
4s3	92.9	89.9		92.9	92.2	93.7	92.9	97.2
4s4	87.2	82.5	90.6		92.6	90.6	96.1	91.8
4s5	91.7	83.7	89.4	89.4		90.2	93.3	91.0
4s6	85.7	93.6	91.7	89.4	88.2		92.9	94.1
4s7	87.0	88.7	90.6	91.8	91.8	92.9		94.1
4s8	90.5	89.9	94.0	87.0	87.0	90.5	91.8	

Figure 4.4 A) Nucleotide sequence and B) amino acid sequence of the novel Vβ17 gene. Sequences are aligned and CDR1 and CDR2 (shaded regions) defined according to the IMGT unique numbering system. Dots identify gaps. C) Percentage nucleotide (top-right) and amino acid (bottom-left) sequence identity of the Vβ17 genes.

A

```
        1           11          21          31          41          51
        .           .           .           .           .           .
VB17s3                                                      CGG AGCTGTGACC

                                          CDR1
       61          71          81          91          101         111
        .           .           .           .           .           .
VB17s3 CTGGAATGTG AACAGGATTT TGATTACGAC TCA....... .......... ....ATGTAC

                                                                  CDR2
      121         131         141         151         161         171
        .           .           .           .           .           .
VB17s3 TGGTACCGAC AGGACCCAGG TCTAGGACTG AGGCGGATCT TCTTCTCACA ATTTGTAAAC

      181         191         201         211         221         231
        .           .           .           .           .           .
VB17s3 ACT....... .....GTTCA GAAAGAAGAC ATAGCT...A AAGGCTACAG CGCCTCTCGT

      241         251         261         271         281         291
        .           .           .           .           .           .
VB17s3 GAA...AAGA AGCCTTTCTT TCCTCTTACC GTGACATCAG CACAGAAGAA TCAGACGGCT

      301         311
        .           .
VB17s3 CTATATCTCT GT
```

B

```
                                          CDR1                      CDR2
        1           11          21          31          41          51
        .           .           .           .           .           .
VB17s3                    GAVT LECEQDFDYD S.......MY WYRQDPGLGL RRIFFSQFVN

       61          71          81          91          101
        .           .           .           .           .
VB17s3 T....VQKED I.AKGYSASR .EKKPFFPLT VTSAQKNQTA LYLC
```

C

Nucleotide sequence identity (%)

Amino acid sequence identity (%)

	17s1	**17s2**	**17s3**
17s1		68.5	96.9
17s2	53.9		69.9
17s3	94.7	58.7	

Figure 4.5 A) Nucleotide sequence and B) amino acid sequence of novel Vβ10 genes. Sequences are aligned and CDR1 and CDR2 (shaded regions) defined according to the IMGT unique numbering system. Identity is shown as dashes and gaps by dots. C) Percentage nucleotide (top-right) and amino acid (bottom-left) sequence identity of the Vβ10genes. Shaded boxes show that Vβ25s1 has <75% nucleotide sequence identity with the Vβ10 genes.

A

```
          1          11          21          31          41          51
          .           .           .           .           .           .
VB10s4                                        GTCAAAG GAAAAGACCA GAAAGCAAAG
VB10s5                                        C------ ---------- --G---C--A
VB10s6                                        ---.... .........- A-GTT-----
VB10s7                                        ------- ------G--- ----------
VB10s8                                        ------- ------G--- --G-------
VB10s9                                        ------- -------A-- --------G-
VB10s10                                       ------- ------G--G ----------
VB10s11                                       ------- ------G--- ----------

                                              CDR1
          61          71          81          91         101         111
          .           .           .           .           .           .
VB10s4    ATGGATTGTG TCCCCATTAA AGGACACGCT TCT....... .......... ....GTTTAC
VB10s5    ---------- ------AA-- -------T-- -A-....... .......... ....------
VB10s6    ---------- -------T-- ------T-T- CA-....... .......... .....--A---
VB10s7    ---------- -------T-- -------AT- AA-....... .......... ....------
VB10s8    ---------- -------A-- ------TAT- AA-....... .......... ....------
VB10s9    ---------- -------A-G ------TAT- AA-....... .......... ....------
VB10s10   ---------- -------A-- ------TAT- AA-....... .......... ....------
VB10s11   ---------- -------A-- ------T-T- AA-....... .......... ....------

                                                                      CDR2
         121         131         141         151         161         171
          .           .           .           .           .           .
VB10s4    TGGTATCTCA AGAAGCTGGA AT---CATTT GAGTTTTTGG TTTATTTGCG GAATGGAAAA
VB10s5    -------G-- G--------- -GAAG----A ---------- ----C--A-A -G---A----
VB10s6    -------A-- ------CA-- -GAAG----- T--------- ---------- ----------
VB10s7    -------G-- ------CA-- -GGAG----- --A------- C---C--AAA ----CA-G-C
VB10s8    -------G-- G--------- -GAAG--C-C -----C--A- ----C--T-A ---GCA-G-C
VB10s9    -------G-- G--------- -GAAG--C-C -----C---- ----C--T-A ---GCA-G-C
VB10s10   -------G-- G--------- -GAAG----- -----C---- ----C--T-A ---GCA-G-C
VB10s11   -------G-- G--------- -GAAG--T-- -----C---- ----C--T-A -G-GCA-G-C

         181         191         201         211         221         231
          .           .           .           .           .           .
VB10s4    ATT....... .....ACAGA AGATACAGCC ATGTTCAAAC AGCGATTTTC AGCTGAGTGC
VB10s5    ---....... .....-T--- ---------- ---------- ---------- --T-------
VB10s6    G--....... ..... T--------- ---------- ---------- ----------
VB10s7    ---....... .....GTG-- --------AA-G-A--TC--- ---------T G---CC---G
VB10s8    ---....... .....GTG-- --GC----AA G--------G ---------- ----------
VB10s9    ---....... .....GTGA- --GC----AA G--------G ---------- ----------
```

```
VB10s10 ---..... .....----- ---------- ---------- ---------T ----------
VB10s11 ---..... .....GTG-- --GC----AA G--------G ------G--- ----------

        241         251         261         271         281         291
         .           .           .           .           .           .
VB10s4  CCC...CAAA ACTCACCCTG CAGCCTGGAA ATCAACTCCA CCGAGGCAGC GGACTCGGCT
VB10s5  ---...---- ---------- ---------- ---------- ---------- ----------
VB10s6  ---...---- ---------- ---------- ---------- ---------- A---------
VB10s7  ---...---- ---------- ---------- ---------- ---------- ----------
VB10s8  -G-...---- ---------- ---------- ---------- ---------T ----G-----
VB10s9  ---...---- ---------- ---------- ---------- ---------- ----G-----
VB10s10 ---...---- ---------- ---------- ---------- ---------T ----------
VB10s11 ---...---- ---------- ---------- ---------- ---------- ----------

        301         311
         .           .
VB10s4  CTGTATTTCT GT
VB10s5  ---------- --
VB10s6  ---------- --
VB10s7  -----C---- --
VB10s8  ---------- --
VB10s9  -----C---- --
VB10s10 ---------- --
VB10s11 -----C---- --
```

B

```
                                      CDR1                          CDR2
        1           11          21          31          41          51
        .           .           .           .           .           .
VB10s4              VKGKDQKAK MDCVPIKGHA S.......VY WYLKKLE.SF EFLVYLWNGK
VB10s5              L-----R-- -----K---S Y.......-- --RR---EAL ------QDE-
VB10s6              V.....SS- ---------V H.......-- --H--P-GA- ----------
VB10s7              ----G---- ---------I N.......-- --R--P-GA- ---A--K-QD
VB10s8              ----G-R-- ---------I N.......-- --RR---EAL -----FQKQD
VB10s9              ----E---R ------R--I N.......-- --RR---EAL -----FQKQD
VB10s10             ----GR--- ---------I N.......-- --RR---EA- -----FQKQD
VB10s11             ----G---- ---------V N.......-- --RR---EA- -----FQEQD

       61          71          81          91          101
        .           .           .           .           .
VB10s4  I....TEDTA MFKQRFSAEC P.QNSPCSLE INSTEAADSA LYFC
VB10s5  I....I---- -------V-- -.-------- ---------- ----
VB10s6  V....-D--- ---------- -.-------- ---------- ----
VB10s7  I....V---E V-Q---L-PW -.-------- ---------- ----
VB10s8  I....V-G-E V--E------ R.-------- ------V-A- ----
VB10s9  I....VKG-E V--E------ -.-------- --------A- ----
VB10s10 I....VKG-E V--E------ -.-------- ------V--- ----
VB10s11 I....V-G-E V--E-C---- -.-------- ---------- ----
```

C

Nucleotide sequence identity (%)

	10s1	10s2	10s3	10s4	10s5	10s6	10s7	10s8	10s9	10s10	10s11	25s1
10s1		93.7	87.4	93.4	88.5	95.6	86.0	83.9	83.9	85.6	86.0	67.5
10s2	91.1		90.3	95.5	90.5	94.7	84.4	83.9	84.8	85.5	86.0	67.1
10s3	80.0	83.3		88.5	90.9	85.2	86.0	87.7	86.8	88.5	88.8	67.1
10s4	91.3	92.5	80.2		90.1	90.1	83.9	84.3	83.9	86.0	86.4	66.3
10s5	80.2	81.5	82.7	81.5		95.6	85.2	87.2	86.4	87.6	88.5	71.0
10s6	96.0	85.5	76.3	86.8	77.6		83.9	82.6	81.7	83.4	83.9	65.3
10s7	80.2	76.5	74.1	77.8	71.6	76.3		87.6	87.6	89.3	89.7	65.9
10s8	71.6	71.6	77.8	72.8	77.8	68.4	77.8		95.8	96.7	95.4	68.7
10s9	71.6	74.1	76.5	72.8	75.3	68.4	76.5	91.3		95.9	95.5	68.3
10s10	74.0	74.1	77.8	75.3	75.3	71.0	80.2	92.6	91.4		97.1	69.8
10s11	77.8	76.5	80.2	77.8	77.8	73.7	81.5	90.1	88.8	92.6		68.3
25s1												

Figure 4.6 A) Nucleotide sequence and B) amino acid sequence of novel Vβ13 and Vβ12 genes. Sequences are aligned and CDR1 and CDR2 (shaded regions) defined according to the IMGT unique numbering system. Identity is shown as dashes and gaps by dots. C) Percentage nucleotide (top-right) and amino acid (bottom-left) sequence identity of the Vβ13 and Vβ12 genes. Shaded boxes indicate >75% nucleotide identity of Vβ12s1 and Vβ12s2 with Vβ13 subfamily members - the value conventionally used to define members of the same subfamily.

A

```
         1           11          21          31          41          51
         .           .           .           .           .           .
VB13s4                                       C CAGGTCATAA GGTCAGGACA GAAACCGACA
VB13s5                                       - ------G-G- --AG------ ---CG-----
VB13s6                                       - ------G-G- --A------T ---CG-----
VB13s7                                       - ------G-G- --AAG----T ---CGT----
VB13s8                                       - ------G-G- --A------- ---CA-----
VB13s9                                       - ------G--- ---------- ---GT-----
VB13s10                                      - ------G--- ---------- ----------
VB13s11                                      - ------G-G- --AG------ --GCG-----
VB13s12                                      - ------G-G- A-A------- ---CGT----
VB13s13                                      - ------G-G- --A------- ---GG-T---
VB13s14                                      - ------G-G- --A------- ---GG-T---
VB13s15                                      - ------G-G- --A------- ---GG-T---
VB13s16                                      - ------G-G- --AG-----T ---CA-----
VB13s17                                      - ------G-G- --A------- ---CAT----
VB13s18                                      - --------G- --A------- ---CGT----
VB13s19                                      - ------G-G- --A------- ---CG-----

VB12s1                                       - ------G--- ---------- ----GT----
VB12s2                                       - ------G--- ---------- ----G-----
```

```
                                                    CDR1
        61          71          81          91          101         111
         .           .           .           .           .           .
VB13s4  CTGAGATGCA CTCAGGATCT GAACCACAAT TAT....... .......... ....ATGTAC
VB13s5  ----A---T- ---------- ---T---G-- ---....... .......... ....------
VB13s6  -A-------- ---------- -------G-- G--....... .......... ....------
VB13s7  ----C---T- ---------- -GGT--TG-C -C-....... .......... ....------
VB13s8  ----C---T- ---------- -GGT--TG-- -GC....... .......... ....------
VB13s9  ---------- ---------- ---------- ---....... .......... ....------
VB13s10 -----G---- ---------- -------GT- ---....... .......... ....------
VB13s11 ----A---T- ---------- -GGT--TG-C CGC....... .......... ....------
VB13s12 --------T- ---------- -T-----G-C AG-....... .......... ....------
VB13s13 T---C---T- -C-------- -GGT--T--C -GC....... .......... ....------
VB13s14 ----C---T- ---------- -G-T--T--C -GC....... .......... ....------
VB13s15 T---C---T- -C-------- -GGT--T--C -GC....... .......... ....------
VB13s16 ----C---T- ---------- --G-TTT-GC ---....... .......... ....------
VB13s17 ---------- ---------- -------G-- ---....... .......... ....------
VB13s18 --------TT ---------- -GGT--T--C -TC....... .......... ....----C-
VB13s19 ----C---T- ---T------ -GGT--TG-C -C-....... .......... ....G-----

VB12s1  ---------- ---------- ---------- ---....... .......... ....------
VB12s2  ---------- ---------- ---------- ---....... .......... ....-C----
```

CDR2

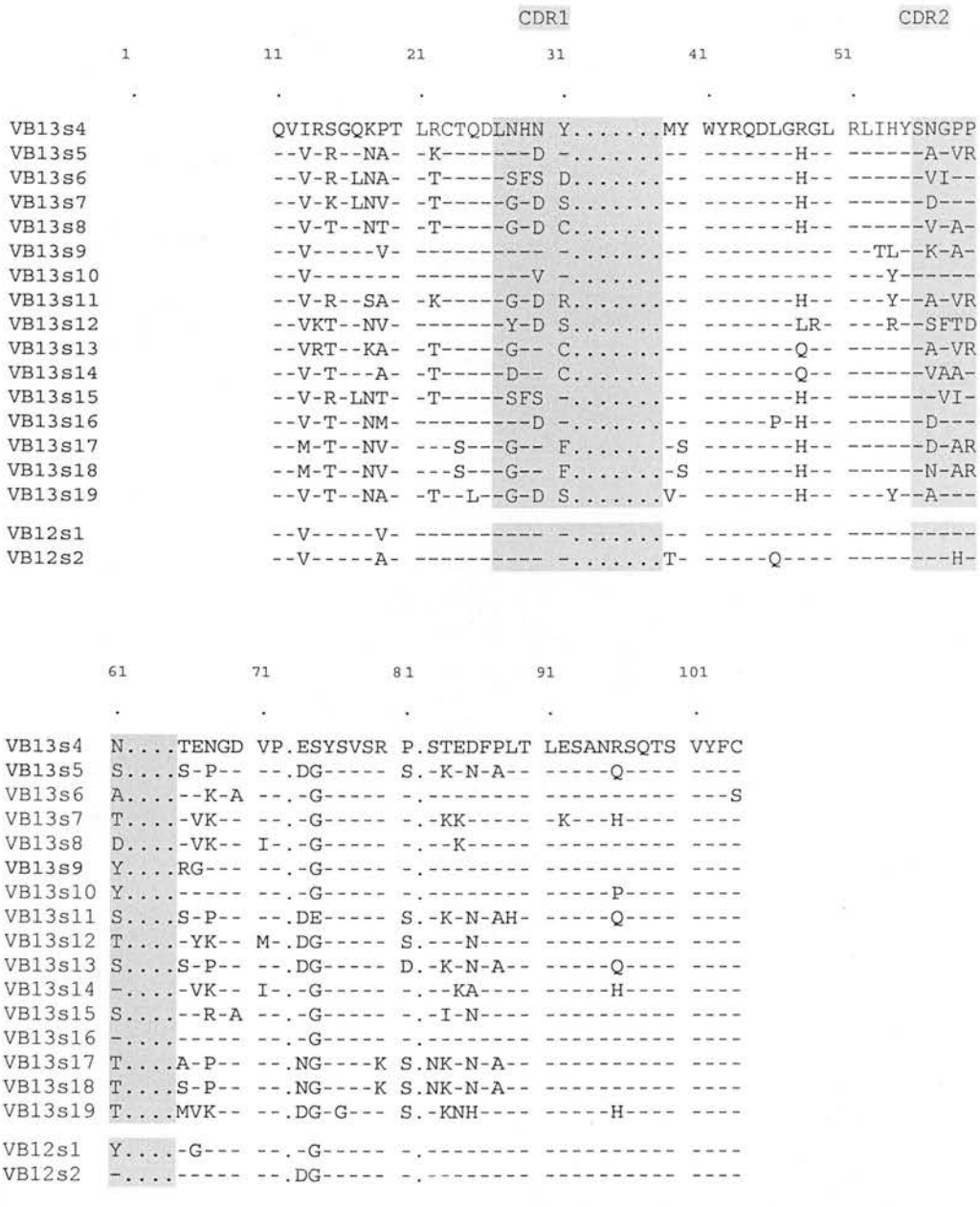
```
          121         131         141         151         161         171
          .           .           .           .           .           .
VB13s4    TGGTACCGAC AAGACCTGGG ACGCGGGCTG AGGCTGATCC ATTACTCAAA TGGCCCTCCC
VB13s5    ---------- ---------- --A------- ---------- ---------GC ---TGTG-GT
VB13s6    ---------- ---------- --A------- ---------- ---------GT -AT------T
VB13s7    ---------- -----T---- --A------- ---------- ---------G- ----------
VB13s8    ---------- -----T---- --AT------ ---------- ---------GT ---TG----T
VB13s9    ---------- ----T----- ---A------ -------CA- T------C-- A--T-----G
VB13s10   ---------- ---------- ---------- ---------T ---------- ----------
VB13s11   ---------- ---------- --A------- ---------T ---------GC ---TGTG-GT
VB13s12   ---------- -------A-- --T-A----- --------T- G---T---TC -TT-A--GA-
VB13s13   ---------- -----T---- --AA------ ---------- ---------GC ----GTG-GT
VB13s14   ---------- -----T---- --AA------ ---------- ---------GT --CTG-----
VB13s15   ---------- -----T---- --AA------ ---------- ---------GT --CTG----T
VB13s16   ---------- ------CA-- --AT------ ---------- ---------GT -AT-------
VB13s17   ---------- ------CA-- --A------- ---------- ---------G- ----------
VB13s18   ---------- ---------- --A------- ---------- ---------G- ---TG-G-G-
VB13s19   ---------- ---------- --A------- ---------T ---------GC ---G------

VB12s1    ---------- ----T----- ---A------ ---------- ---------- ----------
VB12s2    -----T---- ------A--- ---------- ---------- ---------- -----A----
```

```
          181         191         201         211         221         231
          .           .           .           .           .           .
VB13s4    AAC....... .....ACAGA GAACGGAGAT GTGCCC...G AGAGCTACAG CGTCTCCAGA
VB13s5    ---....... .....-GC-- -CCA------ -------...- -CG-G----- ----------
VB13s6    GC-....... .....--G-- ---A----CC -------...- --G------- ----------
VB13s7    -C-....... .....--G-T ---A------ -------...- --G------- ----------
VB13s8    G--....... .....--G-T T--A-----C A-A---...- --G-T----- ----------
VB13s9    T--....... .....AG--G ---------- -----T...- --G------- ----------
VB13s10   T--....... ..... ---------- ---------- ------...- --G------- ----------
VB13s11   TCT....... .....-GC-- -CCA------ -----T...- -CGAG----- ----------
VB13s12   T--....... .....--GT- C--A-----C A----T...- -TG------- ----------
VB13s13   TCT....... .....-GC-- -CCA------ -----T...- -CG-G----- ----------
VB13s14   ---....... .....--G-T T--A-----T A-A---...- --G-T----- ----------
VB13s15   G--....... .....--G-T T--A-----C A-A--T...- --G-T----- ----------
VB13s16   TC-....... .....--G-- --GA----CC -------...- --G------- ----------
VB13s17   ---....... ..... ---------- ---------- --G------- T---------
VB13s18   -CT....... .....-GA-- -CCA-----C --------.A -TG-G----- T-----T-A-
VB13s19   -C-....... .....-TG-T ---A-----C -----T...- -CG-G---G- T---------

VB12s1    T--....... .....----G ---------- -----T...- --G------- ----------
VB12s2    ---....... ..... ---------- ---------- -------...- -CG------- ----------
```

```
        241          251          261          271          281          291
         .            .            .            .            .            .
VB13s4  CCA...AGCA CAGAGGACTT TCCTCTCACA CTGGAGTCTG CCAACCGCTC CCAGACATCT
VB13s5  T--...---- A----A---- -G-------G --------C- ------AG-- ---------C
VB13s6  ---...---- ---------- ---------G ---------- ---------- ----------
VB13s7  ---...---- A-A------- C--------G ---A------ ------A--- ----------
VB13s8  ---...---- --A------- C--------G ---------- ---------- ----------
VB13s9  ---...---- ---------- C--------- ---------- ---------- ----------
VB13s10 ---...---- ---------- C--------- ---------- ------G--- ----------
VB13s11 T--...---- A----A---- -T---A---G --------C- ------AG-- ---------C
VB13s12 T--...---- -----A---- C--------G ---------- ---------- ----------
VB13s13 T--...---- A----A---- -G-------G --------C- ------AG-- ---------C
VB13s14 ---...---- --A---C--- C--------G ---------- ------A--- ----------
VB13s15 ---...---- --A------- C-T------G ---------- ------A--- ----------
VB13s16 ---...---- T----A---- C--------- --------G- ---------- ----------
VB13s17 ---...---- ---------- C--------- ---------- ---------- ----------
VB13s18 T--...-A-- A----A---- -G-A-----G ---------- ---------- ----------
VB13s19 T--...---- A-A-CC---- ---------T --------C- ------A--- ---A-----C

VB12s1  ---...---- ---------- C--------- ---------- ---------- ----------
VB12s2  ---...---- ---------- C--------- ---------- -------T-- ----------
```

```
        301          311
         .            .
VB13s4  GTGTACTTCT GC
VB13s5  --T------- --
VB13s6  ---------A --
VB13s7  ---------- -T
VB13s8  ---------- -T
VB13s9  ---------- --
VB13s10 ---------- --
VB13s11 --T------- -T
VB13s12 ---------- --
VB13s13 --T------- -T
VB13s14 ---------- -T
VB13s15 ---------- -T
VB13s16 ---------- --
VB13s17 ---------- --
VB13s18 ---------- --
VB13s19 ---------- -T

VB12s1  ---------- --
VB12s2  ---------- --
```

B


CDR1
CDR2
1 11 21 31 41 51
VB13s4 QVIRSGQKPT LRCTQDLNHN Y.....MY WYRQDLGRGL RLIHYSNGPP
VB13s5 --V-R--NA- -K-------D -.....-- -------H-- ------A-VR
VB13s6 --V-R-LNA- -T-----SFS D.....-- -------H-- ------VI--
VB13s7 --V-K-LNV- -T-----G-D S.....-- -------H-- ------D---
VB13s8 --V-T--NT- -T-----G-D C.....-- -------H-- ------V-A-
VB13s9 --V-----V- ---------- -.....-- ---------- --TL--K-A-
VB13s10 --V------- ---------V -.....-- ---------- ---Y------
VB13s11 --V-R--SA- -K-----G-D R.....-- -------H-- ---Y--A-VR
VB13s12 --VKT--NV- -------Y-D S.....-- -------LR- ---R--SFTD
VB13s13 --VRT--KA- -T-----G-- C.....-- -------Q-- ------A-VR
VB13s14 --V-T---A- -T-----D-- C.....-- -------Q-- ------VAA-
VB13s15 --V-R-LNT- -T-----SFS -.....-- -------H-- -------VI-
VB13s16 --V-T--NM- ---------D -.....-- -----P-H-- ------D---
VB13s17 --M-T--NV- ---S---G-- F.....-S -------H-- ------D-AR
VB13s18 --M-T--NV- ---S---G-- F.....-S -------H-- ------N-AR
VB13s19 --V-T--NA- -T--L--G-D S.....V- -------H-- ---Y--A---
VB12s1 --V-----V- ---------- -.....-- ---------- ----------
VB12s2 --V-----A- ---------- -.....T- -----Q---- --------H-
61 71 81 91 101
VB13s4 N....TENGD VP.ESYSVSR P.STEDFPLT LESANRSQTS VYFC
VB13s5 S....S-P-- --.DG----- S.-K-N-A-- -----Q---- ----
VB13s6 A....--K-A --.-G----- -.-------- ---------- ---S
VB13s7 T....-VK-- --.-G----- -.-KK----- -K---H---- ----
VB13s8 D....-VK-- I-.-G----- -.--K----- ---------- ----
VB13s9 Y....RG--- --.-G----- -.-------- ---------- ----
VB13s10 Y....----- --.-G----- -.-------- -----P---- ----
VB13s11 S....S-P-- --.DE----- S.-K-N-AH- -----Q---- ----
VB13s12 T....-YK-- M-.DG----- S.---N---- ---------- ----
VB13s13 S....S-P-- --.DG----- D.-K-N-A-- -----Q---- ----
VB13s14 -....-VK-- I-.-G----- -.--KA---- -----H---- ----
VB13s15 S....--R-A --.-G----- -.-I-N---- ---------- ----
VB13s16 -....----- --.-G----- -.-------- ---------- ----
VB13s17 T....A-P-- --.NG----K S.NK-N-A-- ---------- ----
VB13s18 T....S-P-- --.NG----K S.NK-N-A-- ---------- ----
VB13s19 T....MVK-- --.DG-G--- S.-KNH---- -----H---- ----
VB12s1 Y....-G--- --.-G----- -.-------- ---------- ----
VB12s2 -....----- --.DG----- -.-------- ---------- ----

C

Nucleotide sequence identity (%)

Amino acid sequence identity (%)

	13s1	13s2	13s3	13s4	13s5	13s6	13s7	13s8	13s9	13s10	13s11	13s12	13s13	13s14	13s15	13s16	13s17	13s18	13s19	12s1	12s2
13s1		88.4	87.2	84.8	86.1	84.0	90.5	87.7	79.5	84.0	86.5	82.0	86.5	87.7	86.1	84.8	86.1	85.6	97.1	83.6	84.3
13s2	80.9		96.3	81.6	90.2	83.6	85.2	86.1	80.7	81.9	93.4	81.6	88.9	84.0	82.8	83.5	83.6	90.2	86.5	80.7	81.1
13s3	78.6	92.1		81.6	89.3	81.9	84.0	86.9	79.1	80.3	90.2	79.1	88.9	83.2	83.6	84.0	83.2	89.7	84.4	80.6	81.6
13s4	75.3	72.8	74.1		82.4	86.5	86.5	84.8	91.4	96.3	79.5	83.6	78.7	84.8	81.9	85.6	94.3	82.4	81.9	95.9	95.9
13s5	79.0	87.7	83.9	74.1		82.4	84.0	82.8	83.2	82.4	95.1	80.7	93.0	81.6	80.7	81.6	83.9	86.5	85.6	82.4	82.4
13s6	76.5	72.8	71.6	77.8	72.8		88.9	88.1	82.8	86.9	82.6	85.2	80.3	86.1	85.2	93.4	87.7	80.7	84.0	86.9	85.6
13s7	86.4	74.1	72.8	76.5	74.1	80.2		90.9	82.4	86.5	83.6	84.0	84.4	90.2	88.1	87.3	89.3	84.4	89.3	86.9	85.6
13s8	82.7	75.3	75.3	79.0	74.1	81.5	86.4		82.4	84.8	83.6	83.6	84.4	95.5	95.1	86.5	87.3	83.6	86.5	84.8	83.6
13s9	75.3	72.8	72.8	88.9	72.8	76.2	76.5	80.2		90.9	78.8	83.2	80.3	82.4	81.1	81.1	89.3	81.1	77.9	95.5	90.6
13s10	77.8	75.3	72.8	92.6	75.3	77.8	77.8	79.0	90.1		79.1	83.2	78.7	84.0	81.6	85.6	93.8	79.5	81.9	95.5	94.6
13s11	79.0	91.3	82.7	70.4	91.3	69.1	71.6	71.6	69.1	72.8		79.1	93.8	81.6	81.6	79.9	79.5	86.9	86.1	78.3	78.3
13s12	77.8	74.1	70.4	71.6	72.8	70.3	74.1	76.5	74.1	72.8	70.3		78.7	83.2	81.1	83.6	86.5	81.1	81.6	84.4	83.2
13s13	82.7	85.2	85.2	75.3	90.1	71.6	74.1	76.5	74.1	75.3	88.9	72.8		85.2	86.9	80.1	79.5	86.9	85.2	80.3	79.5
13s14	81.5	70.4	70.4	79.0	70.4	77.8	80.2	88.9	79.0	77.8	67.9	72.8	79.0		96.3	84.4	85.6	82.8	85.6	85.2	84.0
13s15	80.2	69.1	69.1	76.5	69.1	76.5	81.5	90.1	77.8	76.5	67.9	70.3	79.0	93.8		82.8	83.2	81.1	84.8	83.6	81.1
13s16	74.1	72.8	75.3	76.5	75.3	90.1	77.8	79.0	74.1	76.5	70.3	69.1	72.8	74.1	71.6		88.9	81.1	82.8	85.6	84.8
13s17	79.0	76.5	75.3	88.9	79.0	81.5	82.7	85.2	85.2	87.6	72.8	76.5	75.3	79.0	76.5	82.7		83.6	84.0	93.8	93.4
13s18	75.3	77.8	79.0	71.6	79.0	66.7	72.8	72.8	71.6	69.1	76.5	71.6	81.5	67.9	67.9	67.9	75.3		83.6	80.7	80.3
13s19	93.8	76.5	71.6	69.1	76.5	71.6	82.7	79.0	69.1	72.8	76.5	74.1	77.8	75.3	75.3	69.1	75.3	70.3		81.6	81.1
12s1	77.8	74.1	75.3	93.8	75.3	79.0	80.2	81.5	95.1	93.8	70.3	75.3	76.5	80.2	79.0	77.8	88.9	72.8	71.6		95.1
12s2	76.5	75.3	75.3	91.4	76.5	76.5	74.1	77.8	87.6	88.9	71.6	72.8	77.8	79.0	76.5	75.3	87.6	70.3	71.6	81.3	

Figure 4.7 A) Nucleotide sequence and B) amino acid sequence of VβXs1, aligned and CDR1 and CDR2 (shaded regions) defined, according to the IMGT unique numbering system.

A

```
      1           11          21          31          41          51
      .           .           .           .           .           .
VBXs1                                                 C CTCGCAGACA GGCTGAAACC

                                                CDR1
      61          71          81          91          101         111
      .           .           .           .           .           .
VBXs1 CTGCGGTGCA TCCTGAAGGA TTCCCAGTAC CCCTGG.... .......... ....ATGAGC

                                                                    CDR2
     121         131         141         151         161         171
      .           .           .           .           .           .
VBXs1 TGGTACCAGC AGGATCTCCG GGGGCAACTA CAGGTGCTGG CCAGTCTGCG GCGTACCGGG

     181         191         201         211         221         231
      .           .           .           .           .           .
VBXs1 GAT....... .....AAGGA AGTCATAAAC CTCCCTGGAG CGAATTACCG GGCCACGCGG

     241         251         261         271         281         291
      .           .           .           .           .           .
VBXs1 GTC...AGTG AGAGCGAGCT GAGCCTACAC GTGGCCAATG TGACC..... .CAGGGCAGA

     301         311
      .           .
VBXs1 ACGCTCTTCT GC
```

B

```
                                        CDR1                        CDR2
      1           11          21          31          41          51
      .           .           .           .           .           .
VBXs1                PRRQAET LRCILKDSQY PW......MS WYQQDLRGQL QVLASLRRTG

      61          71          81          91          101
      .           .           .           .           .
VBXs1 D....KEVIN LPGANYRATR V.SESELSLH VANVT..QGR TLFC
```

were represented by only one transcript (e.g. Vβ13s12). This would indicate that although this strategy enabled a rapid and focused examination of a large number of unique TCRβ transcripts bearing Vβ genes of specific subfamilies and the identification of several new members, comprehensive characterisation of the membership of these subfamilies had probably still not been achieved. A similar but smaller scale experiment examined 32 Vβ4 clones derived from animal 641 and identified 2 novel genes - Vβ4s7 and Vβ4s8 (Figure 4.3).

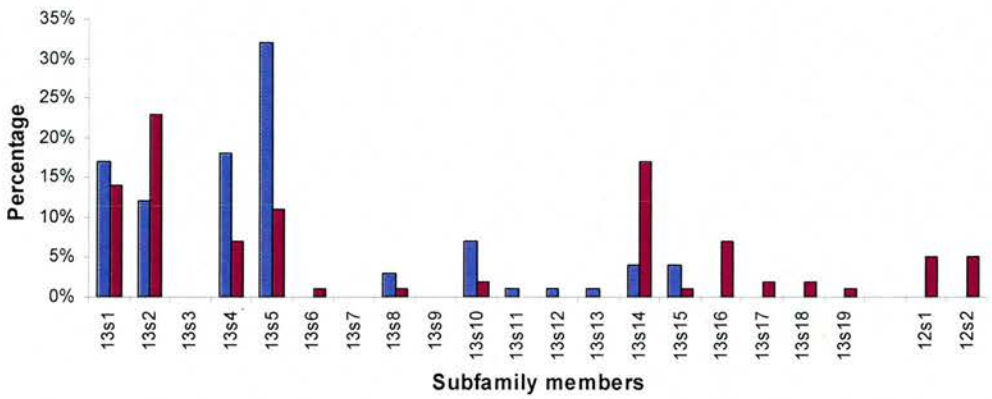


Figure 4.8 Percentage of TCRβ transcript sub-clones generated from animals 605 (blue) and 641 (maroon) using the 5'primer VB13ext, that expressed the different Vβ13 and 12 subfamily members.

4.3.3 Additional Vβ genes amplified by Vβ10 and Vβ13 subfamily-specific primers

In addition to Vβ13 genes, both Vβ12s1 and Vβ12s2 genes were amplified from the cDNA of animal 641 by the 5'primer VB13ext (Figure 4.8). As stated before (section 4.3.1), both Vβ12 members were initially identified following amplification with Vβ13 subfamily-specific primers and so this was not unexpected. A single TCRβ chain transcript amplified by VB10ext was found to have highest nucleotide sequence identity to human Vβ25s1 (77.9%, compared to 68.0% nucleotide identity with human Vβ10s1) and had <75% nucleotide homology with all of the bovine Vβ10 genes (Figure 4.5c); accordingly the gene was named Vβ25s1. The nucleotide and amino acid sequence data for Vβ25s1 is given in Figure 4.9.

Figure 4.9 A) Nucleotide sequence and B) amino acid sequence of Vβ25s1, aligned and CDR1 and CDR2 (shaded regions) defined, according to the IMGT unique numbering system.

A

```
        1           11          21          31          41          51
        .           .           .           .           .           .
VB25s1                                      GTTCAAG GGAAAGGACT GGAAGTGAAA

                                            CDR1
       61          71          81          91         101         111
        .           .           .           .           .           .
VB25s1 ATGTATTGTG TTCCCAAAAA AGGACATATT TAT..... ..........  ....GTGTTC

                                                                    CDR2
      121         131         141         151         161         171
        .           .           .           .           .           .
VB25s1 TGGTATCAGC AGATCCTGAC CAAAGAGTTC AAGTTCTTGA TTTCCTTCCA AAATGATAAA

      181         191         201         211         221         231
        .           .           .           .           .           .
VB25s1 GTC.......  .....TTGGA TGATAAAGAG ATGCCC...A AGAGATTTTC AGCTGAGTGC

      241         251         261         271         281         291
        .           .           .           .           .           .
VB25s1 CCC...TCAA ACTCACCTTG TAGTCTGAAG ATCCAGCCTA CGGAGCCGCA GGACTCAGCT

      301         311
        .           .
VB25s1 ATGTATTTCT GT
```

B

```
                                    CDR1                        CDR2
        1           11          21          31          41          51
        .           .           .           .           .           .
VB25s1              VQGKGLEVK MYCVPKKGHI Y.......VF WYQQILTKEF KFLISFQNDK

       61          71          81          91         101
        .           .           .           .           .
VB25s1 V....LDDKE MP.KRFSAEC P.SNSPCSLK IQPTEPQDSA MYFC
```

4.3.4 Demonstration of Vβ1, Vβ2, and Vβ13 subfamily expansion in genomic DNA

An alternative approach to examining the extent of membership of bovine Vβ subfamilies was to examine un-rearranged genomic DNA using Southern blot analysis with subfamily-specific probes.

Genomic DNA from a primary fibroblast cell-line was digested with four enzymes (*Hind*III, *Ssp*I, *Xba*I and *Sph*I), selected on the basis that identified Vβ genes belonging to the subfamilies being examined did not have the respective restriction sites. ^{32}P-labelled probes specific for Vβ1, Vβ2, Vβ13 and Vβ14 (representing a putative single member subfamily) were generated; due to the high nucleotide similarity between members of the Vβ1 and 5 subfamilies and the Vβ12 and 13 subfamilies (Figure 4.1c and Figure 4.6c) it was assumed that the Vβ1 probe would bind to both Vβ1 and Vβ5 genes and the Vβ13 probe would bind both Vβ12 and Vβ13 genes.

Hybridization with the Vβ14 probe revealed a single band (5.0-7.9kb in size) in preparations of DNA digested with *Hind*III, *Ssp*I and *Sph*I (Figure 4.10a) - consistent with Vβ14 being a single member subfamily. The results of hybridisation with the Vβ2 probe were dependent on the enzyme that had been employed for digestion. In the *Xba*I digested-DNA at least eight bands of varying intensity, ranging in size from 3.0 - 10.0 Kb, were discernible (Figure 4.10b, lane 3), whilst in the *Hind*III digested-DNA only two bands of 4.1 and 4.4 Kb were detected (Figure 4.10b lane 1) and in each of the *Ssp*I and *Sph*I digested preparations there were 4 bands (Figure 4.10b, lanes 2 and 4). The marked variation in the intensity of bands detected with the Vβ2 probe suggests that some of the bands may represent superimposition of two or more genes.

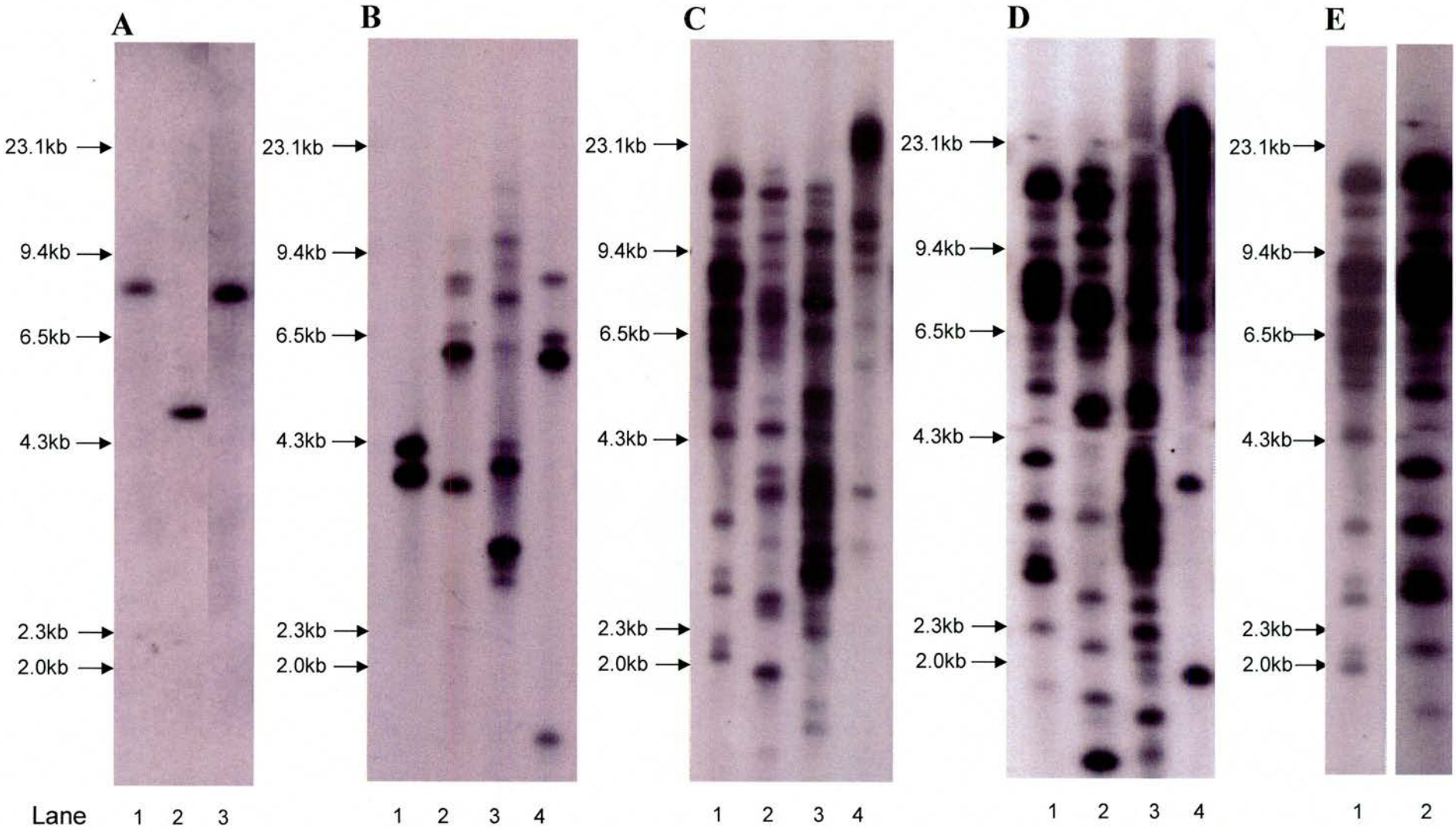


Figure 4.10 Southern blot analysis of bovine digested genomic DNA with TCRβV probes. Genomic DNA digested with by HindIII (lane 1), SspI(lane 2), XbaI (lane 3) and SphI (lane 4), hybridised with probes for (A) Vβ14, (B) Vβ2, (C) Vβ1 and (D) Vβ13. (E) Southern blot of bovine genomic DNA digested with HindIII with probes for Vβ1 (lane 1) and Vβ13 (lane 2)

Both the Vβ1 and Vβ13 probes gave a highly complex banding pattern (Figure 4.10c and d), with a minimum of 15 bands ranging in size from 1.0 - 14.0Kb detected with three out of the four enzymes (*Hind*III, *Ssp*I and *Xba*I). Again variation in intensity of the bands suggested that some of the bands represented superimposed genes. Notably, there was a high degree of similarity in the banding pattern of the Vβ1 and Vβ13 probes (Figure 4.10e), suggesting that some of the Vβ1 and Vβ13 genes may be located on the same DNA fragments.

Thus, the results of the southern blots corroborate the cDNA data, demonstrating that several bovine TCRβV subfamilies contain large numbers of members. However, (i) potential superimposition of multiple DNA fragments sharing restriction sites in some bands, (ii) the possibility that some of the larger bands may represent DNA fragments containing multiple Vβ genes belonging to the same subfamily and (iii) likely cross-reactivity of subfamily specific probes prevents precise enumeration of the membership of these subfamilies.

4.3.5 Analysis of bovine Jβ genes from rearranged TCRβ chain transcripts.

Analysis of the initial dataset of 173 unique TCRβ chains revealed the existence of 16 different Jβ gene sequences, including all 13 that had been described in previous studies (Houston, 1997; Tanaka *et al.*, 1990); no additional Jβ genes were identified in the TCRβ transcripts subsequently generated. All of the Jβ gene sequences identified were present in more than one unique TCRβ rearrangement, and thus sequence variation is assumed to represent genuine germline sequence variation rather than experimental error due to mis-incorporation of nucleotides. As Jβ nomenclature is based on genomic position, which was unavailable at the time of analysis, the Jβ genes were given a temporary numeric nomenclature (J1-16), with the three novel Jβ genes identified being J14-16.

Each of the bovine Jβ sequences shared high nucleotide (70-93%) and amino acid (60-100%) sequence identity to human orthologues (Table 4.2). There appeared to be two bovine orthologues to each of four human Jβ genes in the second genomic cluster (Jβ2s1, 2s2, 2s5 and 2s7). The two bovine orthologues of human Jβ2s5 (J6 and J8) differed from each other by a single nucleotide substitution and similarly J7 and J15, both orthologues of human Jβ2s1, differed by only two base pairs; therefore these pairs of sequences may represent allelic variants of the same Jβ gene. However, the two bovine orthologues of human Jβ2s2 - J1 and J10 - demonstrated significant sequence divergence (12/30bp difference - only 60% nucleotide identity), as did the two bovine orthologues of human Jβ2s7 - J3 and J13 - (6/30bp difference – only 80% nucleotide identity); suggesting that these J genes are products of different gene loci rather than allelic variants.

The frequencies with which the Jβ genes were expressed were highly variable, with one of the novel genes, J16, being only identified in two of the approximately 400 unique TCRβ rearrangements analysed, whilst others such as J2 were utilized in ~20% of TCRβ transcripts. However, this skewed frequency of Jβ usage may reflect the biased nature in which the majority of these transcripts were generated (i.e. from antigen specific T-cell clones or Vβ subfamily-specific PCR).

Bovine Jβ gene	Human orthologue	Bovine and human J sequences	Nucleotide identity (%)	Amino acid identity (%)
J1	2s2	TTTGGAGGGGGCACTCGGCTGTTGGTACTA -------AA---T--A-----ACC-----G	73	70
J2	1s2	TTCGGCCCAGGGACCAAGCTGACGGTTGTA -----TT-G-------G-T-G--C------	77	80
J3	2s7	TTCGGCCCAGGCACCAAGCTCACGGTCTTA -----G--G-------G----------AC-	83	80
J4	2s3	TTCGGCGCGGGCACCCGGCTGACGGTGCTA --T---C-A-------------A-----C	83	90
J5	2s4	TTCGGCGCGGGCACCCGGCTGTCGGTGCTA --------C--G--------C--A-----G	83	100
J6	2s5	TTCGGGCCGGGCACGCGGCTCCTGGTGCTA --------A--------------------C	93	100
J7	2s1	TTCGGGCCAGGGACCAGGCTCATTGTGCTA --------------AC------CC------	87	90
J8	2s5	TTCGGGCCGGGCACGCGGCTCCAGGTGCTA ----------------------T------C	93	90
J9	1s5	TTTGGACACGGGACCTGGCTCGCCGTCCTG -----TG-T-----TC-A---T--A----A	70	60
J10	2s2	TTTGGAGCTGGTTCCAAGCTGACTGTGCTG -------AA--C--T-G------C--A---	77	80
J11	1s4	TTCGGCAACGGGACGAAGCTTTCCGTCTTG --T----GT--A--CC----C--T------	73	80
J12	2s6	TTCGGGGCCGGCAGCTGGCTGACGGTGGTG --------------A-------C---C--	90	80
J13	2s7	TTTGGCGCCGGCACCAGGCTCACGGTCACA --C--GC-G--------------------	87	90
J14	1s1	TTTGGAAAAGGCACCAGGCTCACGGTTGTA ------C----------A-----A------	90	90
J15	2s1	TTCGGGCCAGGGACCAGGCTCTCTGTGCTA --------------AC------AC------	87	90
J16	1s6	TTTGGAATCGGCACCAGGCTCATCGTGACA -----A-AT--G----------CT------	80	80

Table 4.2 The 16 bovine Jβ genes identified from TCRβ chain transcripts. The human orthologue for each gene (column 2) was assumed to be that showing highest nucleotide (column 4) and amino acid (column 5) sequence identity. Nucleotide alignment of the bovine Jβ genes (upper) with their human orthologues (Lopes *et al.*) is shown in column 3, with identical base pairs shown as dashes. The three novel bovine Jβ genes identified in this study are J14-16.

4.4 Discussion

The results of analysing over 400 unique bovine TCRβ chain rearrangements presented in this chapter constitute a significant advance towards the complete characterisation of the bovine TCRβ gene repertoire. Analysis of ~150 bovine TCRβ rearrangements in previous studies had established a repertoire of 46 Vβ genes in 17 subfamilies and 13 Jβ genes (Buitkamp *et al.*, 1993 Russell and MacHugh, unpublished; Housset *et al.*, 1997; Houston, 1997; Tanaka *et al.*, 1990). In this study an additional 38 novel bovine Vβ genes, including members of 4 novel subfamilies (Vβ5, 12, 25 and X), and 3 novel Jβ genes were identified, bringing the currently known bovine TCRβ gene repertoire to 84 Vβ genes distributed over 21 subfamilies, and 16 Jβ genes.

In the absence of Vβ gene germline sequence data, a major challenge to Vβ gene classification is the difficulty in defining sequences sharing high nucleotide identity as allelic variants or products of unique but closely related subfamily members. Based on observations made by Arden *et al.* (1995a,b) on human and murine Vβ sequences, it has been convention to assume that sequences differing by less than 5 nucleotides (or sharing ≥98% nucleotide identity) are allelic variants. However, there are demonstrable exceptions to this convention - for example, human Vβ8s1 and Vβ8s2 differ by only 5 nucleotides (Siu *et al.*, 1986) whereas the two alleles of murine Vβ10s1 differ by 9 nucleotides (Smith *et al.*, 1990). For the analysis completed in this study the convention was strictly adhered to. Some of the novel Vβ genes displayed levels of nucleotide identity with other subfamily members approaching 98%, but such high levels of nucleotide identity between different members was not atypical for previously identified genes within these subfamilies (Figures 4.1-4.6 parts c). This may reflect recent Vβ duplication events with minimal subsequent divergence, or gene conversion such as that seen with human Vβ13s2a and Vβ13s2b (Glusman *et al.*, 2001). Until the full genomic sequence of the bovine TCRB locus is known classification of closely related Vβ sequences will retain an element of uncertainty.

Analysis of Vβ sequences derived from amplified TCRβ chain mRNA transcripts is further complicated because sequence variation may reflect artefacts generated by (i) modification of the 3' end nucleotide sequence of Vβ genes during somatic recombination or (ii) mis-incorporation of nucleotides during PCR, replication of plasmid sub-clones or sequencing. In this study the sequence 3' to the conserved cysteine at position 104 was excluded from analysis, removing somatic recombination as a potential source of non-germline sequence variation. Although it is not possible to eliminate sequence variation from mis-incorporation of nucleotides, most of the novel Vβ genes identified (27 out of 38) were present in multiple transcripts, suggesting they reflect genuine germline sequence variation rather than technical artefacts. Of the 11 novel Vβ genes identified from single transcripts, matching sequences for 5 (Vβ1s12, Vβ10s5, Vβ10s6, Vβ13s12 and Vβ25s1) were later identified in the available bovine genome data (see Chapter 5), and it is anticipated that future studies will produce evidence corroborating the sequences of the remaining 6 (Vβ2s8, Vβ2s9, Vβ4s7, Vβ4s8, Vβ13s11, Vβ13s13).

By convention Vβ genes of non-human/non-murine species have been designated subfamily numbers corresponding to the human Vβ with which they display highest nucleotide identity (Baron *et al.*, 2001; Li *et al.*, 1990; Schrenzel *et al.*, 1994; Tanaka *et al.*, 1990). Adherence to this convention resulted in 3 genes being ascribed to the novel bovine subfamilies Vβ5 and Vβ12 (section 4.3.1). However, by the criterion of members of the same Vβ subfamily sharing >75% nucleotide identity (Arden *et al.*, 1995a; Kazatchkine, 1995) both Vβ12 genes could be considered to be members of the bovine Vβ13 subfamily (Figure 4.6c) and similarly the single Vβ5 gene shares >75% nucleotide identity with 7 out of the 13 bovine Vβ1 members (Figure 4.1c). As such, the utilisation of this convention may be artificially sub-dividing bovine subfamilies and demonstrates the difficulties that can be encountered by extrapolating conventions concerning TCR gene nomenclature across species.

The number of genes constituting the known bovine Vβ repertoire (84 Vβ genes) now exceeds that of humans (65 Vβ genes - Rowen *et al.*, 1996), making it the largest Vβ repertoire characterised to date. The bovine Vβ repertoire demonstrates a

substantial level of gene duplication, with 11 of the 21 bovine Vβ subfamilies identified containing multiple members (Table 4.3). More strikingly, results from this study have added Vβ4, 10 and 13 to Vβ1 and 2 as bovine subfamilies which have undergone extensive duplication, each having ≥8 members. Sixty of the 84 genes (over 70%) in the Vβ repertoire belong to one of these 5 subfamilies. Despite intensive sequencing of a large number of $V\beta10^+$ and $V\beta13^+$ TCRβ transcripts, results suggest that this analysis has not been exhaustive and further members of these subfamilies remain to be identified (section 4.3.2). The use of subfamily-specific primers to generate Vβ sequences for analysis will have undoubtedly favoured identification of multiple members of the selected subfamilies. It is possible that similar investigations of other subfamilies would extend the known membership of some of them, although an analysis of 32 Vβ7 transcripts amplified using the subfamily-specific primer failed to identify any new members of this subfamily (data not shown).

The large membership of the Vβ1, 2 and 13 subfamilies was evident in analysis of restriction enzyme digested genomic DNA by southern blot, where subfamily-specific probes hybridised with multiple bands. Marked variation in the intensity of the bands was notable in several of the lanes in Figure 4.10 (e.g. Figure 4.10b lane 3), suggestive that some of the bands represented superimposed copies of subfamily members. This most likely reflects the fact that recently duplicated Vβ genes in the same subfamily typically have conserved flanking non-coding sequence and therefore retain some of the same restriction sites (Glusman *et al.*, 2001). Such a phenomenon has been observed for some murine TCRαV genes (Clark *et al.*, 1995) and may also account for a previous underestimate of the number of bovine Vβ2 genes made from southern blot data of *PstI* digested genomic DNA (Tanaka *et al.*, 1990).

Comparisons of Vβ sequences from different mammalian species provide convincing evidence that ancient duplications of an ancestral Vβ gene predating mammalian radiation established many of the different Vβ subfamilies

Vβ subfamily	Bovine	Human	Porcine	Ovine	Caprine
1	13	1	Yes	5	Yes
2	9	1	Yes	1	Yes
3	1	1		1	Yes
4	8	1	Yes	1	Yes
5	1	8	Yes		
6	3	9	Multigenic	1	Yes
7	2	3	Multigenic	1	Yes
8	2	5	Yes	1	Yes
9	1	2	Yes		Yes
10	11	1	Yes	1	Yes
11		1	Yes		
12	2	3	Yes		
13	19	9	Yes	2	Yes
14	1	1	Yes		
15	2	1		1	Yes
16	1	1			
17	3	1	Yes	1	Yes
18		1			
19		1			
20	1	1	Yes		Yes
21		3	Yes		
22		1	Multigenic		
23		1			
24	1	1	Yes	1	Yes
25	1	1			Yes
26		1			
27		1			
28	1	1		1	
29		1			
30		2			Yes
X	1		Yes		
Total	84	65		18	

Table 4.3 The bovine, human, porcine, ovine and caprine Vβ gene repertoires. Details of the number of members in porcine and caprine subfamilies is limited and so YES is used to indicate subfamilies in which Vβ genes have been identified. Porcine subfamilies Vβ6, 7 and 22 are known to be multigenic (Baron *et al.* 2001; Butler *et al.* 2005).

(Baron *et al.*, 2001; Clark *et al.*, 1995; Houston and Morrison, 1999). Consistent with this view of Vβ evolution, all of the 21 bovine subfamilies now identified have >70% nucleotide sequence identity to human orthologues, with the exception of VβX, which is an orthologue of murine Vβ2. The human orthologue of this gene appears to have been deleted (Glusman *et al.*, 2001), but an orthologue has been identified in the pig (Baron *et al.*, 2001) which, like cattle, is an artiodactyl (even-toed ungulate) species. Given the conservation of Vβ subfamilies across the mammalian spectrum it is possible that bovine orthologues to at least some of the other 10 human Vβ subfamilies will be found with further investigation.

Due to duplications and deletions of the TCRB locus that have occurred subsequent to mammalian radiation there is variation in the germline Vβ gene repertoires of individual species (Table 4.3 - Baron *et al.*, 2001; Bosc and Lefranc, 2000; Folch and Lefranc, 2000b; Halsey *et al.*, 1999). As the retention of duplications in particular Vβ subfamilies indicates that they are likely to convey an evolutionary advantage for the survival of a species, similarities and differences in which subfamilies have undergone duplication in various species are of interest. In contrast to the mouse, in which only 2 of the 31 subfamilies have multiple members, 9 of the 30 human subfamilies are multi-membered (section 1.2.3.2) as are 11 of the 21 bovine subfamilies so far identified. Some of the expanded subfamilies are common to human and cattle (e.g. Vβ6, Vβ7 and Vβ8) whereas others differ (e.g.Vβ2, Vβ4 and Vβ10 are duplicated in cattle but not human).

It is intriguing to note that in both human and cattle the Vβ13 subfamily has been subjected to extensive amplification, as have the Vβ5 subfamily in humans and Vβ1 subfamily in cattle. As commented on above, phylogeny of human Vβ1 and 5 genes (Arden *et al.*, 1995a), and nucleotide sequence identity of bovine Vβ1 and 5 genes (section 4.3.1) demonstrate that these two subfamilies are intimately related. In humans, expansion of the Vβ5 and Vβ13 subfamilies is largely due to multiple duplication of a cassette containing Vβ5 and Vβ13 elements (Rowen *et al.*, 1996; Su and Nei, 2001). Substantial similarity in the banding patterns produced by the Vβ1 and Vβ13 probes in southern blots of digested bovine genomic DNA (Figure 4.10e)

suggests that in cattle there also may have been extensive duplication of a cassette containing a Vβ1 and Vβ13 gene. As it has been estimated that the human cassette containing the Vβ5 and Vβ13 genes underwent duplication in only the last 32 million years (Su and Nei, 2001), which is after human/bovine divergence 65-80 MYA (Li *et al.*, 1990), the expansion of human Vβ5 and 13 and bovine Vβ1 and 13 subfamilies may represent equivalent but distinct evolutionary developments (see also section 5.4).

Data on the Vβ repertoires of three other artiodactyl species - sheep, goat and pig - have been published. The pig is the only other artiodactyl species whose Vβ repertoire has been subject to extensive analysis (Baron *et al.*, 2001; Butler *et al.*, 2005). Members of 19 subfamilies have been identified, including 3 subfamilies that are absent from the bovine repertoire (Vβ11, Vβ21 and Vβ22). Baron *et al.* (2001) identified 2 members in both Vβ6 and Vβ22 subfamilies and Butler *et al.* (2005) described several different Vβ7 sequences. However, comprehensive details of the numbers of Vβ genes in the different porcine subfamilies are awaiting publication (Butler *et al.*, 2005). A limited study in sheep, based on analysis of 34 TCRβ sequences, identified 18 Vβ genes in 13 subfamilies (Halsey *et al.*, 1999) and in goats, analysis of 55 sequences identified members of 19 subfamilies (Obexer-Ruff *et al.*, 1998). With the exception of caprine subfamily Vβ30, all of the subfamilies identified in these two ruminant species are represented in the bovine repertoire. Notably, the only subfamilies identified in sheep as having duplications were Vβ1 and Vβ13 (Halsey *et al.*, 1999).

Sixteen bovine Jβ sequences have been identified, which again exceeds the number of Jβ genes identified in any other species to date. All of the Jβ genes shared ≥70% to human orthologues, which is similar to the levels of homology identified between orthologous human and porcine Jβ genes (Baron *et al.*, 2001).

Sequencing of the region including the full repertoire of Dβ, Jβ and Cβ genes has been achieved for the human (Rowen *et al.*, 1996), murine (GenBank accession number AE000665), rat (Williams *et al.*, 1991) and porcine (GenBank accession

number AB079894) TCRB loci. In all cases the Dβ, Jβ and Cβ genes are located in tandemly duplicated DJC clusters, each containing a single Dβ gene, several (six or seven) Jβ genes and a single constant (Cβ) gene and there is a high level of inter-species conservation of the Dβ and Jβ genes in both sequence and organisation (Baron *et al.*, 2001; Schrenzel *et al.*, 1994; Williams *et al.*, 1991).

Intriguingly, several of the human Jβ genes in the second DJC cluster appear to have two bovine orthologues. The orthologues of human Jβ2s5 and Jβ2s1 show minimal sequence variation and may represent allelic variants of the same gene, but the orthologues of Jβ2s2 and Jβ2s7 display significant sequence divergence and are likely to represent different gene loci. Based on recent preliminary analysis of genomic sequence data, it has been suggested that the bovine TCRB locus contains a third constant region gene (Conrad *et al.*, 2002). A duplication event incorporating the Jβ2 cluster and a Cβ gene could account for both the identification of two bovine orthologues for several Jβ2 genes and the presence of a third Cβ gene. In the sheep, 10 Jβ genes have been identified, including a Jβ2s2 orthologue and a Jβ gene which was unclassified but the amino acid sequence of which is identical to bovine J10, also an orthologue of Jβ2s2 (Halsey *et al.*, 1999). The existence of two Jβ2s2 orthologues in the sheep is suggestive that this extra duplication event occurred prior to ovine/bovine divergence 22 MYA (Hassanin and Douzery, 2003) and may be a feature of other ruminant species.

In conclusion, the results presented in this study have expanded the known repertoire of bovine TCRβ genes, which appears to be characterised by a high level of duplication of both Vβ and Jβ genes. Notably, several subfamilies, Vβ1, 2, 4, 10 and 13, have undergone prodigious duplication. Identification of these new bovine TCRβ genes will be helpful in the development and refinement of methods to study T-cell responses in cattle. However, it is likely that additional bovine TCRβ genes remain to be identified.

Part of the southern blot data presented in this chapter contributed to the publication Houston *et al.* (2005) which has been reproduced at the end of this thesis.

5 Chapter 5 – The Bovine TCRB locus

5.1 Introduction

T-cells can be segregated into two distinct lineages that express either an αβTCR (αβT-cells) or a γδTCR (γδT-cells). The α,β,γ and δ polypeptide chains that form these heterodimeric receptors each consist of a constant domain that anchors the receptor to the T-cell membrane and a variable domain that participates in antigen recognition. The variable domains are encoded by exons formed during T-cell ontogeny by somatic recombination of variable (V), diversity (D - β and δ chains only) and joining (J) gene segments.

In humans and mice, the TCR genes are located in three chromosomal loci - TCRB (TCRβ chain genes), TCRG (TCRγ chain genes) and TCRA/D (TCRα and δ chain genes). The complete sequencing of the human and murine TCR loci has been pivotal in providing a better understanding of TCR genes in these species (Glusman *et al.*, 2001). Firstly, it has enabled a definitive catalogue of the TCR gene segments to be obtained, thus enabling functional TCR studies to proceed without potential complications arising from incomplete knowledge of the TCR gene repertoire. Secondly, availability of genomic sequences of the TCR loci has allowed identification and systematic analysis of the promoter and enhancer elements that regulate TCR gene expression (Anderson *et al.*, 1988; Krimpenfort *et al.*, 1988; Mathieu *et al.*, 2000; Rowen *et al.*, 1996; Ryu *et al.*, 2004) and the effect that recombination signal sequences (RSS) can have on utilization of particular TCR gene segments (Probst *et al.*, 2004; Wu *et al.*, 2003). Thirdly, the ability to study the genomic organisation of the TCR loci has provided insights and new perspectives on the evolutionary history of this important group of genes (Glusman *et al.*, 2001; Su and Nei, 2001).

The TCRB locus is situated on chromosome 7q34 in humans and chromosome 6A-C in mice (Barker *et al.*, 1984; Caccia *et al.*, 1984; Isobe *et al.*, 1985). The genomic organisation of the locus is conserved between humans and mice; the 5' end contains

an array of V genes (65 in humans and 35 in mice), with the D, J and C genes arranged into tandem clusters containing a single Dβ gene, six or seven Jβ genes and a single Cβ gene at the 3'end of the locus. A solitary Vβ gene with an inverted transcriptional orientation to the other TCRβ genes lies at the extreme 3' end of the locus. Within the Vβ array the relative order of human and murine orthologues displays a striking level of conservation (Lai *et al.*, 1988; Rowen *et al.*, 1996), and similarly the arrangement of orthologous genes in the DJC clusters is highly conserved. The conservation of synteny also extends to non-TCRβ genes in, and adjacent to, the TCRB locus. In both species multiple trypsinogen genes flank the 5' and 3' ends of the array of Vβ genes, the ephrin type-b receptor 6 precursor (EPHB6) gene is situated at the extreme 3'end of the locus and a dopamine-β-hydroxylase-like (DBH-like) gene is located immediately 5' to the locus (Glusman *et al.*, 2001).

In humans, inter-chromosomal translocation has produced a cluster of Vβ pseudogenes located on chromosome 9 (Glusman *et al.*, 2001; Robinson *et al.*, 1993; Rowen *et al.*, 1996). These Vβ genes are unable to undergo recombination because no Dβ, Jβ or Cβ genes were incorporated in this inter-chromosomal translocation, and are described as 'orphan' genes.

Sequence data of TCR loci in other mammals such as cat, dog, sheep, pig, and cow is currently being generated through efforts to sequence the entire genomes of these species (http://www.ncbi.nih.gov/Genomes). The bovine genome project, an international effort headed by Dr. Richard Gibbs, of the Baylor College of Medicine Human Genome Sequencing Centre (Houston, Texas, USA), began sequencing and assembling the bovine genome in December 2003. The project is focused on a whole genome shotgun (WGS) analysis of a female Hereford named L1 Dominette 01449. In October 2004 a first draft 3.3-fold WGS assembly was released, followed by a second draft 6.2-fold WGS assembly in March 2005. Genomic sequencing of the entire bovine TCRG loci (GenBank accession numbers AY644517 and AY644518) has already proved useful in detailing the genomic organisation of the TCRγ genes and, in combination with cDNA analysis, in defining the full catalogue of TCRγ gene segments in cattle (Herzig *et al.*, 2006a). The bovine genome has also be used

recently to examine the TCRδ gene repertoire and organisation (Herzig *et al.*, 2006b).

Extensive cDNA analysis has achieved characterisation of a large number of bovine TCRβ gene segments (this study; Buitkamp *et al.*, 1993; Houston 1997; Houston and Morrison 1999; Tanake *et al.*, 1990; Russell and MacHugh unpublished data) but it is unlikely that the full complement of bovine TCRβ genes has been described. In contrast to the TCRG locus, only a small portion of the TCRB locus, incorporating part of the DJC region, has been deposited (GenBank accession number AF453325). It was therefore decided to examine the current bovine genome assembly to further characterise the repertoire of TCRβ genes and the TCRB locus and compare the repertoire of expressed TCRβ genes described in the previous chapter with that present in the genomic assembly. Although the genomic sequence of the TCRB locus is incomplete, the results presented in this chapter reveal that (i) there is a massive germline Vβ gene repertoire in the bovine, (ii) there are three DJC clusters, as had been suggested previously (Conrad *et al.*, 2002) and (iii) there is conserved synteny in the organisation of the bovine and human TCRB loci.

5.2 Materials and methods

5.2.1 Sequence data

Partial sequences of bovine Vβ and Jβ genes identified from cDNA analysis of rearranged TCRβ transcripts were derived from data collated in Chapter 4. Available sequences of germline bovine Dβ, Jβ and Cβ genes were derived from the annotated sequence of part of the Dβ, Jβ, Cβ region of the bovine TCRB locus (GenBank accession number AF453325 - Conrad *et al.*, 2002). Human germline sequence data for Vβ, Dβ, Jβ and Cβ genes was extracted from the annotated TCRB locus sequence deposited in the GenBank database (U66059-U66061 - Rowen *et al.*, 1996).

5.2.2 Bovine genome analysis

The bovine genome assembly and WGS 'trace' sequence archive were accessed through the National Centre for Biotechnology Information (NCBI) BLAST facility (http://www.ncbi.nlm.nih.gov/BLAST). Sequences were compared to genomic data using the BLASTN algorithm under default conditions. Sequences were deemed to show significant homology if the 'expected' (E) value was below 0.001. The current bovine chromosome map and annotation of genomic scaffolds were accessed through the NCBI Map Viewer (http://www.ncbi.nlm.nih.gov/mapview).

Where possible, the entire Vβ gene from initiation codon of the leader (L) exon to the end of the 3'flanking 23bp spacer recombination signal sequence (23bp-RSS) was identified and extracted. Leader exons were identified from the limited data on bovine L exon available from deposited sequences (section 2.5.1) or similarity to orthologous human Vβ gene L exon sequence. The Vβ intron was defined by identification of conventional GT/AG splice sites in conjunction with knowledge of the predicted L exon and variable (V) exon boundaries. RSS sequences were defined according to the conserved consensus heptamer-spacer-nonamer characteristics (Hesse *et al.*, 1989; Ramsden *et al.*, 1994). Jβ gene sequence data extracted included the 5' flanking 12bp-RSS and the 3' flanking guanine (G) residue that is donated to

the Cβ gene during splicing. Dβ gene sequence included both 5' flanking 23-bp RSS and 3' flanking 12bp-RSS sequences.

5.2.3 Sequence analysis

Some general details about sequence analysis are given in section 2.5. Using the IMGT/V-QUEST software programme (Giudicelli *et al.*, 2004 - http://imgt.cines.fr) the human Vβ gene with the highest nucleotide similarity to each genomic Vβ gene sequence was identified and the subfamily of the bovine Vβ genes thus determined.

Percentage nucleotide identities were calculated using the Lasergene software (DNAstar, Madison, WI). Comparison of Vβ sequences included only the sequence encoding the mature Vβ polypeptide.

Based on experimental data from human and murine systems, which have provided information on sequence features that are essential for TCRβ gene segments to be functional, the following criteria were evaluated in determining the potential functional integrity of genomic bovine TCRβ genes: (i) The presence of an initiation codon in the predicted L exon of Vβ genes, (ii) maintenance of an open reading frame in the L and V exons of Vβ genes in the post-splicing configuration and in Jβ genes in the same reading frame as encodes the conserved Phe-Gly-X-Gly motif, (iii) the presence of sequence encoding the conserved residues at positions 23 (cysteine), 41 (tryptophan), 42 (tyrosine) and 104 (cysteine) of the mature Vβ polypeptide, (iv) the presence of correct GT/AG splice sites at the end of the L exon and beginning of V exons of Vβ genes, (v) the presence of a RSS compatible with somatic recombination at the 3'flank of Vβ genes (23-bp spacer), 5'flank of Jβ genes (12-bp spacer) and the 3' (23-bp spacer) and 5' (12-bp spacer) flank of Dβ genes. The consensus sequence of the human and murine RSS is CACAGTG-12/23bp-ACAAAAACC although there is variation in the heptamer and the nonamer sequence (Ramsden *et al.*, 1994). *In vitro* studies have indicated that the initial three CAC bases of the heptamer are critical for recombination (Akamatsu *et al.*, 1994; Hesse *et al.*, 1989) as is a non-A base at the either of the two most distal (from the

spacer) positions in the nonamer (Akamatsu *et al.*, 1994); there are conflicting results concerning the requirement for three consecutive As in the nonamer core (Akamatsu *et al.*, 1994; Hesse *et al.*, 1989) and so in this study they have not been considered necessary for recombination. RSSs with spacers varying in length by +/-1bp have been shown to be capable of undergoing recombination, whereas spacers that differ in length by >1bp from the usual 12 or 23bp are unable to effect recombination (Akamatsu *et al.*, 1994; Hesse *et al.*, 1989).

5.2.4 Generation and sequencing of Vβ11$^+$ TCRβ chain transcripts

5.2.4.1 Vβ11$^+$ subfamily-specific PCR amplification of cDNA

Naïve PBMC from animal 641 were isolated (section 2.3.1) and cDNA generated as described in sections 2.4.1-2.4.3. cDNA was amplified using a Vβ11-specific 5'primer (CCCAAGATACTGTGTGATAGGGAT) and the Cβ specific 3'primer BCext (Table 2.3). The reaction used 20 pmol of each primer, 1 unit BIOTAQ (5units/μl Bioline, London, UK), 4 μl SM-0005 buffer (ABgene, Epsom, Surrey, UK – Appendix C.7), 2 μl cDNA, and nuclease-free water to give a final volume of 40 μl. The programme used was as follows – 5 min at 94°C, 5 cycles of (1 min at 94°C, 1 min at 60°C, 1 min at 72°C), 25 cycles of (30 s at 94°C, 1 min at 60°C, 1 min at 72°C) and a final extension period of 5 min at 72°C.

5.2.4.2 Cloning and sequencing of Vβ11$^+$ TCRβ chain transcripts

Following confirmation of PCR products of the expected size by agarose gel electrophoresis (section 2.4.5), the products were purified using the Wizard PCR Preps DNA Purification System (section 2.4.6), sub-cloned into plasmid pGEM-T Easy vectors and used to transform High Efficiency Competent JM109 cells (section 2.4.7). The presence of inserts of the anticipated size in selected sub-clones was verified by PCR as described in section 2.4.7.1. The resulting PCR products were sequenced (section 2.4.9.1-2.4.9.2) and the sequences analysed as described in section 4.2.2.

5.3 Results

5.3.1 Characterization of the Vβ gene repertoire in the current bovine genome assembly

5.3.1.1 Identification of 123 bovine Vβ genes

Sequences within the bovine genome showing significant sequence homology to the 84 known bovine Vβ genes defined from analyses of cDNA (Chapter 4), and/or the 65 human TCRβV genes in the TCRB locus (Rowen *et al.*, 1996) were identified using the BLASTN algorithm: 138 sequences were found on 36 scaffolds/contigs. The subfamily to which each sequence belonged was determined and, to avoid confusion with the nomenclature used for the cDNA data, each sequence was assigned a name composed of the subfamily and an alphabetic suffix (e.g. VB13a). As commented on previously (section 4.3.1), there was high nucleotide sequence similarity between genes belonging to Vβ1 and 5 subfamilies and between genes belonging to the Vβ12 and Vβ13 subfamilies (data not shown).

It was not possible to assign the severely mutated genes named Untitled 1 and Untitled 2 to subfamilies. Both showed highest nucleotide identity with human Vβ12 genes (70% nucleotide identity between Untitled 1 and human Vβ12s2, 74% nucleotide identity between Untitled 2 and human Vβ12s1) but neither had >75% nucleotide identity with any of the members of the bovine Vβ12 subfamily (data not shown) and thus could not be described as Vβ12 genes.

Comparisons of percentage nucleotide sequence identities of genes within the same subfamilies (data not shown) demonstrated that several genes shared ≥98% nucleotide identity. Some of these segments were defined as distinct genes on the basis that they were either (i) located on the same contig (e.g. VB1v and VB1w) or (ii) were adjacent to different Vβ genes at equivalent positions in their respective contigs (e.g. VB1p and VB1aa). Fifteen genes sharing >98% nucleotide sequence similarity to other sequences could not be verified as either allelic variants or distinct genes and were therefore left unclassified and not included in further analysis

(Appendix G). Exclusion of these unclassified genes gave a total of 123 Vβ genes, located on 26 scaffolds/contigs, in the current bovine genome assembly (Table 5.1).

Genes belonging to several new subfamilies were found, including single members of subfamilies Vβ11 (80.1% nucleotide identity to human Vβ11s1), Vβ21 (highest identity with human Vβ21s1 - 75.9%), Vβ23 (75.6% nucleotide identity to human Vβ23s1) and three members of Vβ18 (76.6%, 76.3% and 76.4% nucleotide identity to human Vβ18s1).

The most striking feature of the repertoire of Vβ genes identified from the genome was the massive numbers of genes belonging to subfamilies Vβ1 (29) and Vβ13 (36). Extensive duplication of a cassette containing tandemly arranged Vβ1/13 genes accounts for the majority of this expansion, with several of the large scaffolds/contigs containing up to five consecutive repeats of this cassette (e.g. NW_931372.1/Bt4_WGA842_2 and NW_937104.1/BtUn_WGA12501_2 - Table 5.1).

Table 5.1 Bovine Vβ genes in the current bovine genome assembly (overleaf). The name, sequence orientation and length of the scaffolds/contigs on which Vβ sequences have been identified are shown in column 1. The location on the scaffolds/contigs of the Vβ sequences from initiation codon of the leader exon to the end of the RSS sequence is shown in column 2; the lengths of the sequences are detailed in parentheses. The nine Vβ sequences (VB2d, VB10c, 10q, 10u, VB13z, 13ab, VB17d, VB18c and Untitled 1) for which genomic sequence was incomplete are marked by (*). The name and functional status for each Vβ sequence is shown in columns 3 and 4; genes were classified as either functional (Func), pseudogenes (P), open-reading frame pseudogenes (ORF), orphan pseudogenes (Orphan) or unknown, as described in section 5.3.1.3. Vβ gene sequences derived from cDNA that displayed ≥98% nucleotide identity to genomic sequences were considered to represent matches (section 5.3.1.4) and are shown in column 4. In column 5 the number of matching base pairs is shown as a fraction of the total as well as a percentage.

Scaffold/Contig	Location (size in bp)	Name	Functionality	Matching cDNA sequence	Nucleotide identity between genomic and cDNA sequences
NW_931372.1/	1,942-1,429 (514)	VB1a	Func		
Bt4_WGA842_2	4,108 – 3,638 (471)	VB13a	P		
(Rev) 397,044	7,249-6,781 (469)	VB13b	P		
	9,478-8,966 (513)	VB1b	P		
	11,711-11,251 (461)	VB13c	P		
	13,421-12,906 (516)	VB1c	Func	Vβ1s11	257/259 (99%)
	16,375-15,907 (469)	VB13d	ORF		
	20,308-19,795 (514)	VB1d	Func	Vβ1s5	265/267 (99%)
	24,183-23,715 (469)	VB13e	Func	Vβ13s4	240/244 (98%)
	127,658-127,142 (517)	VB1e	Func	Vβ1s8	271/271 (100%)
	134,709-133,563 (517)	VB5a	P		
	144,067-143,573 (495)	VB7a	Func	Vβ7s1	270/273 (98%)
	148,348-147,849 (500)	VB9a	Func	Vβ9s1	214/214 (100%)
	224,480-223,797 (663)	VBxa	Func	Vβxs1	234/234 (100%)
NW_931373.1/	2,007-2,4174 (468)	VB13f	ORF		
Bt4_WGA843_2	3,730-4,238 (509)	VB1f	P		
(For) 117,860	5,662-6,123 (462)	VB13g	ORF	Vβ13s14	241/244 (98%)
	7,860-8,357 (498)	VB1g	Func		
	13,438-13,886 (449)	VB12a	P		
	15,444-15,957 (514)	VB5b	P		
	17,335-17,804 (470)	VB13h	Func	Vβ13s5	241/244 (98%)
	22,731-23,199 (469)	VB13i	Func		
	24,934-25,431 (498)	VB1h	Func		
	26,786-27,255 (470)	VB13j	Func	Vβ13s18	240/244 (98%)
	29,030-29,545 (516)	VB1i	Func		
	31,525-31,993 (469)	VB13k	Func	Vβ13s12	243/244 (99%)
	34,432-34,916 (485)	VB6a	Func	Vβ6s1	239/240 (99%)
	38,570-38,972 (403)*	Untitled 1	P		
	49,381-49,887 (507)	VB5c	Func	Vβ5s1	167/169 (98%)
	54,673-55,065 (623)	VB6b	Func	Vβ6s2	268/268 (100%)
	59,624-60,117 (494)	VB23a	P		
	66,135-66,601 (467)	Untitled 2	P		
	71,941-72,436 (496)	VB21a	P		
	78,923-79,404 (482)	VB8a	Func	Vβ8s1	282/282 (100%)
	93,732-94,016 (481)	VB8b	Func	Vβ8s2	275/276 (99%)
	96,082-96,549 (468)	VB16a	Func	Vβ16s1	267/270 (98%)

Table 5.1 Bovine Vβ genes in the current bovine genome assembly (1 of 5 pages). Notes on page 154.

Scaffold/Contig	Location (size in bp)	Name	Functionality	Matching cDNA sequence	Nucleotide identity between genomic and cDNA sequences
NW_931373.1/	98,381-98,884 (504)	VB24a	Func	Vβ24s1	271/273 (99%)
Bt4_WGA843_2	103,112-103,610 (489)	VB25a	P	Vβ25s1	238/240 (99%)
(For) 117,860. Cont.	114,007-114,671 (665)	VB18a	P		
NW_931380.1/	94,139-94,838 (700)	VB20a	Func	Vβ20s1	272/273 (99%)
Bt4_WGA850_2	154,796-154,162 (635)	VB4a	Func	Vβ4s6	249/253 (98%)
(Rev) 175,468	157,271-156,634 (638)	VB4b	Func	Vβ4s4	279/282 (98%)
	159,755-159,122 (633)	VB4c	Func	Vβ4s3	252/253 (99%)
	162,835-162,156 (680)	VB4d	P		
	168,549-167,838 (712)	VB3a	Func	Vβ3s1	273/273 (100%)
	173,797-173,289 (509)	VB14a	Func	Vβ14s1	272/273 (99%)
NW_931534.1/	72,447-72,916 (470)	VB13l	ORPHAN	Vβ13s3	262/267 (98%)
Bt5_WGA944_2					
(For) 437,222					
NW_935493.1/	5,820-6,318 (499)	VB10a	P		
BtUn_WGA10890_2	10,098-10,588 (491)	VB10b	Func	Vβ10s8	239/243 (98%)
(For) 45,197	19,101-19,593 (492)*	VB10c	Unknown	Vβ10s3	225/225 (100%)
	23,860-24,366 (507)	VB17s2a	Func	Vβ17s2	270/270 (100%)
	32,873-33,375 (503)	VB15a	Func		
	38,632-39,141 (510)	VB11a	Func	Vβ11s1	249/251 (99%)
	40,332-40,850 (519)	VB28a	Func	Vβ28s1	265/266 (99%)
NW_936628.1/	1,796-2,264 (469)	VB13m	P		
BtUn_WGA12025_2	3,961–4,476 (516)	VB1j	Func	Vβ1s13	169/169 (100%)
(For) 37,581	8,419-8,932 (514)	VB1k	Func	Vβ1s3	272/273 (99%)
	9,958-10,426 (469)	VB13n	ORF		
	11,793-12,300 (508)	VB1l	P		
	13,776-14,245 (470)	VB13o	P		
	15,174-15,642 (470)	VB13p	ORF		
	17,561-18,059 (499)	VB1m	Func	Vβ1s2	271/273 (99%)
	22,269-22,782 (514)	VB1n	P		
	24,133-24,585 (453)	VB13q	P		
	28,340-28,834 (495)	VB7b	Func	Vβ7s1	270/273 (98%)
NW_937104.1/	347-816 (470)	VB13r	Func	Vβ13s10	243/244 (99%)
BtUn_WGA12501_2	4,077-4,589 (513)	VB1o	Func		
(For) 34,940	8,007-8,4375 (469)	VB13s	ORF		
	10,279-10,794 (516)	VB1p	Func	Vβ1s10	224/227 (98%)
	11,987-12,456 (470)	VB13t	P		

Table 5.1 (continued) Bovine Vβ genes in the current bovine genome assembly (2 of 5 pages).

Notes on page 154.

Scaffold/Contig	Location (size in bp)	Name	Functionality	Matching cDNA sequence	Nucleotide identity between genomic and cDNA sequences
NW_937104.1/	14,441-14,959 (519)	VB1q	Func		
BtUn_WGA12501_2	16,135-16,604 (470)	VB13u	P		
(For) 34,940 Cont.	18,294-18,809 (516)	VB1r	Func		
	22,689-23,202 (514)	VB1s	Func		
	24,548-25,016 (469)	VB13v	Func		
	26,913-27,381 (469)	VB13w	Func	Vβ13s16	243/244 (99%)
	28,763-29,271 (509)	VB1t	P		
	30,697-31,165 (469)	VB13x	Func	Vβ13s7	244/244 (100%)
	32,098-32,567 (470)	VB13y	Func	Vβ13s19	242/244 (99%)
NW_937259.1/	255-1 (255)*	VB13z	P		
BtUn_WGA12656_2	3,531-3,062 (470)	VB13aa	ORF		
(Rev) 34,012	5,393-4,896 (498)	VB1u	Func	Vβ1s12	216/218 (99%)
	7,474-7,185 (290)*	VB13ab	Unknown	Vβ13s18	117/117 (100%)
	9,348-8,834 (515)	VB1v	P	Vβ1s9	269/270 (99%)
	11,993-11,524 (470)	VB13ac	ORF		
	15,676-15,161 (516)	VB1w	Func	Vβ1s9	269/270 (99%)
	17,894-17,425 (470)	VB13ad	ORF		
	25,787-25,291 (497)	VB1x	P		
	27,993-27,525 (469)	VB13ae	Func	Vβ13s8	243/244 (99%)
	29,900-29,390 (511)	VB5d	P		
	31,605-31,137 (469)	VB13af	P		
NW_939383.1/	857-86 (880)	VB2a	Func	Vβ2s2	279/281 (99%)
BtUn_WGA14780_2	3,834-3,319 (516)	VB17b	P	Vβ17s1	269/273 (98%)
(Rev) 24,070	5,948-5,750 (670)	VB18b	P		
	11,275-10,748 (528)	VB10d	P		
	22,034-21,566 (469)	VB10e	Func	Vβ10s4	239/240 (99%)
NW_939582.1/	7,455-7,942 (488)	VB10f	P		
BtUn_WGA14979_2	13,260-13,749 (490)	VB10g	Func	Vβ10s7	241/243 (99%)
(For) 23,220	16,749-17,249 (501)	VB10h	P		
	21,957-22,444 (488)	VB10i	Func	Vβ10s10	240/243 (98%)
				Vβ10s11	239/243 (98%)
NW_940681.1/	14,460-13,608 (853)	VB2b	Func	Vβ2s5	280/282 (99%)
BtUn_WGA16078_2	18,998-18,935 (215)*	VB18c	P		
(Rev) 19,041					

Table 5.1 (continued) Bovine Vβ genes in the current bovine genome assembly (3 of 5 pages).

Notes on page 154.

Scaffold/Contig	Location (size in bp)	Name	Functionality	Matching cDNA sequence	Nucleotide identity between genomic and cDNA sequences
NW_941023.1/ BtUn_WGA16420_2 (For) 17,866	3,188-3,704 (517) 8,087-8,555 (469) 11,791-12,304 (514) 17,029-17,544 (516)	VB1y VB13ag VB1z VB1aa	Func Func Func Func	Vβ1s8 Vβ13s4 Vβ1s5 Vβ1s10	270/271 (99%) 241/244 (98%) 266/267 (99%) 224/227 (98%)
NW_942019.1/ BtUn_WGA17416_2 (For) 14,735	1,816–2,869 (1,053)	VB2c	ORF		
NW_942641.1/ BtUn_WGA18038_2 (For) 12,789	5,737-6,191 (455)*	VB2d	Unknown	Vβ2s6	281/282 (99%)
NW_943204.1/ BtUn_WGA18601_2 (Rev) 11,302	2,427-1,912 (516) 10,551-10,062 (490)	VB17c VB10m	Func Func	Vβ17s1 Vβ10s11	268/273 (98%) 240/243 (98%)
NW_944159.1/ BtUn-WGA19556_2 (Rev) 8,982	3,273-2,795 (479)	VB10n	P	Vβ10s6	229/230 (99%)
NW_944919.1/ BtUn_WGA20316_2 (Rev) 7,403	3,629-3,132 (498)	VB10o	ORF		
NW_945250.1/ BtUn_WGA20647_2 (Rev) 6,781	3,460-2,964 (497) 5,670-5,209 (462)	VB1ab VB13aj	P ORF	 Vβ13s14	 240/244 (98%)
NW_945323.1/ BtUn_WGA20720_2 (Rev) 6,654	1,765-1,280 (486) 6,654-6,412 (243)*	VB10p VB10q	Func Unknown		
NW_946773.1/ BtUn_WGA22170_2 (For) 4,821	719-1,481 (856)	VB2f	Func	Vβ2s7	280/282 (99%)
NW_947228.1/ BtUn_WGA22625_2 (Rev) 4,370	4,320-3,827 (494)	VB10r	P		
NW_94104.1/ BtUn_WGA24501_2 (Rev) 3,201	1,725-1,115 (611)*	VB17d	Unknown		

Table 5.1 (continued) Bovine Vβ genes in the current bovine genome assembly (4 of 5 pages).

Notes on page 154.

Scaffold/Contig	Location (size in bp)	Name	Functionality	Matching cDNA sequence	Nucleotide identity between genomic and cDNA sequences
NW_951386.1/ BtUn_WGA26783_2 (For) 2,409	692-1,160 (469)	VB13ak	Func		
NW_954128.1/ BtUn_WGA29525_2 (Rev) 2,058	1,950-1,436 (515)	VB1ac	ORF		
NW_963357.1/ BtUn_WGA38754_2 (For) 1,517	1-339 (339)*	VB10u	Unknown	Vβ10s1	274/276 (99%)
NW_937259.1/ BtUn_WGA63389_2 (Rev) 1,025	1,018-549 (470)	VB13am	Func	Vβ13s1	272/272 (100%)

Table 5.1 (continued) Bovine Vβ genes in the current bovine genome assembly (5 of 5 pages). Notes on page 153.

5.3.1.2 All but one of the mapped Vβ genes are located on chromosome 4

The chromosomal locations of the contigs containing Vβ genes were undetermined in the available annotation of the genome sequence data, but 3 large scaffolds - NW_931372.1/Bt4_WGA842_2, NW_931373.1/Bt4_WGA843_2, NW_931380.1/Bt4_WGA850_2 - which between them contained 46 Vβ genes were mapped to chromosome 4.

One gene, VB13l, was located on a scaffold - NW_931534.1/Bt5_WGA944_2 - mapped to chromosome 5, inferring that in cattle, as in humans (Robinson *et al.*, 1993; Rowen *et al.*, 1996), there has been inter-chromosomal translocation of Vβ gene(s). Consequently, VB13l has been tentatively designated an orphan pseudogene as no Dβ, Jβ or Cβ have been identified on chromosome 5 (section 5.3.2-4).

5.3.1.3 Functional integrity of the Vβ genes identified in the genome

The sequence of the entire gene, from the initiation codon of the L exon to the end of the recombination signal sequence (RSS), was available for 114 of the 123 Vβ genes identified in the genome sequence data (the incomplete sequences were: VB2d, VB10c, 10q, 10u, VB13z, 13ab, VB17d, VB18c and Untitled 1). The leader sequence, splice sites, translated Vβ region and the RSS were all analysed for each of these 114 sequences to assess their predicted functional status.

Vβ genes can be assigned to one of four functional classes: (i) Functional, (ii) pseudogenes (P), which lack an initiation codon or contain a premature stop codon, usually due to frameshifts within the coding regions, (iii) open-reading frame pseudogenes (ORFs), which are non-functional due to defects in splice sites or RSSs or loss of conserved residues vital for protein function, and (iv) orphan genes, which are non-functional due to absence of adjacent D, J and C gene segments on the same chromosome.

Using these criteria, only 65 of the 114 intact Vβ genes identified in the genome are predicted to be functional. The remaining 49 genes are predicted to encode 35 pseudogenes, 13 ORFs pseudogenes and a single orphan gene. Of the 9 genes for which incomplete sequence data was obtained 3 (Untitled 1, VB 18c and 13z) would be predicted to encode pseudogenes whilst the functional status of the remaining 6 can not be determined (Table 5.1 and Table 5.3). A summary of the results obtained from this analysis are presented in Appendix H.

5.3.1.4 Comparison of cDNA and genomic Vβ gene sequences

Of the 84 Vβ genes identified from TCRβ transcripts, 55 (65%) showed ≥98% nucleotide identity to a genomic Vβ sequence and were considered to be matches; corresponding cDNA and genomic Vβ sequences are shown in Table 5.1. An additional 3 Vβ genes identified from cDNA (Vβ2s1, 2s3, 13s2) had matches in 'trace' sequences from the WGS analysis that have not been integrated into the genomic assembly (Table 5.2). Of the 26 Vβ genes identified from cDNA that have

no matches in the current genome assembly or 'trace' sequences, 21 come from the large multi-member subfamilies (Vβ1s1, 1s4, 1s6, 1s7; Vβ2s4, 2s8, 2s9; Vβ4s1, 4s2, 4s5, 4s7, 4s8; Vβ10s2, 10s5, 10s9; Vβ12s1, 12s2; Vβ13s6, 13s9, 13s11, 13s13), with Vβ6s3, Vβ7s2, Vβ15s1, 15s2 and Vβ17s3 also unmatched to any genomic sequence.

Vβ gene	WGS Trace	Location	Matches	Sequence identity
2s1	382505214	538-257	1-282	281/282 (99%)
2s3	491324250	897-616	1-282	280/282 (99%)
13s2	648423488	802-536	1-267	267/267 (100%)

Table 5.2 Vβ gene sequences identified from cDNA that display ≥98% nucleotide identity with WGS 'trace' sequences.

Twelve of functional Vβ genes identified within the genome (VB1a, 1g, 1k, 1i, 1o, 1q, 1r, 1s: VB10p; VB13i, 13v and 13ak) had no corresponding cDNA matches. Notably, except for VB11a, all of these genes belong to the large subfamilies Vβ1, 10 and 13. A Vβ11 subfamily-specific 5'primer was designed and used together with the BCext primer to amplify Vβ11 segments from cDNA prepared from PBMC. Analysis of the sequences of 36 cDNA clones derived from the PCR product, confirmed the presence of a Vβ11 subfamily member (Vβ11s1), the nucleotide and amino acid sequences of which are given in Figure 5.1. The nucleotide sequence identity between the genomic and transcript consensus sequences for Vβ11s1 was 99% (Table 5.1).

Two of the Vβ genes identified from rearranged TCRβ transcripts matched genomic genes predicted to be unable to undergo somatic recombination. Vβ13s14 matched VB13g and aj (Table 5.1), both of which had defects in the nonamer sequence of the RSS which were considered to prohibit recombination (Appendix H). Vβ13s3 matched VB13l (Table 5.1), which has been designated as an orphan pseudogene due to its location on chromosome 5 (section 5.3.1.2). A possible explanation for

Figure 5.1 A) Nucleotide sequence and B) amino acid sequence of Vβ11s1, aligned and CDR1 and CDR2 (shaded regions) defined, according to the IMGT unique numbering system.

A

```
        1           11          21          31          41          51
        .           .           .           .           .           .
VB11s1                          CAAGATAC TGTGTGATAG GGATGGGAAA GAAGATTACT

                                                  CDR1
       61          71          81          91         101         111
        .           .           .           .           .           .
VB11s1 CTAGAATGTT CTCAAACTAT GGGCCTTGAC AAC..... .......... ....ATGTAC

                                                                  CDR2
      121         131         141         151         161         171
        .           .           .           .           .           .
VB11s1 TGGTATCAGC AAGACCCAGG AATGGAACTG CAGCTGATCC ATTATTCATA CGGGGTTAAT

      181         191         201         211         221         231
        .           .           .           .           .           .
VB11s1 ACC..... .....ACGGA GAAAGGAGAG CGCTCG...T CTAGGTCGAC TGTCTCCAGA

      241         251         261         271         281         291
        .           .           .           .           .           .
VB11s1 ACA...AGCA AAGAGCACTT TCCCCTGACG CTAGAGACCG CCAACTCCTC ACAGACGTCT
      301         311
        .           .
VB11s1 AGCTACTTCT GC
```

B

```
                                    CDR1                          CDR2
        1           11          21          31          41          51
        .           .           .           .           .           .
VB11s1          RY CVIGMGKKIT LECSQTMGLD N.......MY WYQQDPGMEL QLIHYSYGVN

       61          71          81          91         101
        .           .           .           .           .
VB11s1 T....TEKGE RS.SRSTVSR T.SKEHFPLT LETANSSQTS SYFC
```

these anomalous matches is that there are highly homologous but functional and as yet unidentified Vβ genes within the bovine TCRB locus. As discussed in section 5.3.1, there are several examples of distinct genes sharing ≥98% nucleotide sequence identity and, as already exemplified with Vβ1s9 and Vβ17s1, it is possible for Vβ genes identified from cDNA to match multiple genomic sequences (VB1v/VB1w and VB17b/VB17c respectively), of which only one is functional (VB1w and VB17c respectively). Alternative explanations include (i) errors in analysis of the genomic data, e.g. in sequencing or chromosomal mapping, (ii) mis-classification of genes with RSS defects that may not absolutely preclude recombination in cattle and (iii) existence of functional allelic variants of either VB13g or aj which lack the RSS defect.

Amalgamating data from the cDNA and genomic analysis gives a total repertoire of 152 bovine Vβ genes belonging to 25 different subfamilies (Table 5.3). Thirteen subfamilies have multiple members, with subfamilies Vβ1, 2, 4, 10, and 13 representing large subfamilies with ≥ 9 members each. Most notable is the exceptional numbers of genes within the Vβ10 (18 members), Vβ1 (33 members) and Vβ13 (42 members) subfamilies.

Table 5.3 The repertoire of bovine Vβ genes identified from combined genomic and cDNA analysis (overleaf). Numbers in parentheses identify the number of genes identified in both cDNA and genomic DNA. Vβ sequences identified only from cDNA are not classified into functional and non-functional groups as functional status can not always be determined from cDNA data.

Vβ sub-family	Genomic Data					cDNA Data	Total
	Functional	Pseudo-gene	ORF	Orphan	Unknown		
1	20	8	1			13 (9)	33
2	3		1		1	9 (4)	11
3	1					1 (1)	1
4	3	1				8 (3)	9
5	1	3				1 (1)	4
6	2					3 (2)	3
7	2					2 (1)	3
8	2					2 (2)	2
9	1					1 (1)	1
10	6	6	1		3	11 (8)	18
11	1					1 (1)	1
12		1				2 (0)	3
13	14	10	10	1	1	19 (14)	42
14	1					1 (1)	1
15	1					2 (0)	3
16	1					1 (1)	1
17	2	1			1	3 (2)	4
18		3					3
19							
20	1					1 (1)	1
21		1					1
22							
23		1					1
24	1					1 (1)	1
25		1				1 (1)	1
26							
27							
28	1					1 (1)	1
29							
30							
X	1					1 (1)	1
Untitled		2					2
Total	65	38	13	1	6	85 (56)	152

Table 5.3 The repertoire of bovine Vβ genes identified from combined genomic and cDNA analysis. Notes on previous page.

5.3.2 Analysis of the Jβ gene repertoire in the current bovine genome assembly.

Seventeen Jβ segments were identified in the bovine genome, based on significant sequence homology to (i) the partial bovine Jβ gene sequences identified from analysis of TCRβ transcripts, (ii) the 13 available bovine germline Jβ sequences and/or, (iii) human germline Jβ gene sequences. These sequences were all located on scaffold NW_931380.1/Bt4_WGA850_2 and organised into three clusters at ≈131,000-133,000 (5 sequences), ≈122,000-123,000 (5 sequences) and ≈112,000-114,000 (7 sequences). As the scaffold sequence was in reverse orientation, these were referred to as bovine Jβ1, Jβ2 and Jβ3 clusters respectively. Assuming that all of the bovine Jβ genes were identified, the conventional nomenclature system, whereby Jβ genes are named according to genomic position was adopted (Kazatchkine, 1995).

From the genomic data it was possible to identify the entire germline sequence of each Jβ gene, including the 12bp-RSS (Table 5.4). Comparison of the human and bovine germline Jβ sequences allowed orthologues to be identified (Appendix I, Table 5.4) and, as had been indicated from cDNA analysis (section 4.4), the presence of two bovine clusters (Jβ2 and Jβ3) orthologous to the human Jβ2 cluster was confirmed; this will be discussed in more detail in section 5.3.5. Although none of the bovine clusters contained orthologues to all of the genes in the orthologous human Jβ cluster, there appears to be conservation of the genomic organisation of the Jβ genes within clusters (Figure 5.2).

Each of the Jβ sequences derived from cDNA analysis had exact identity with a single sequence in the genome except for J10 and J2. The genomic segments Jβ2s2 and Jβ3s2 both showed exact identity with the 30bp of J10 on which classification of the cDNAs was based, but their germline sequence differed 3' to the conserved Phe-Gly-X-Gly motif. Retrospective analysis suggests that both Jβ2s2 and Jβ3s2 genes were represented among the TCRβ transcripts that had been defined as expressing J10 (data not shown). The genomic segment Jβ1s2 differed from J2 by a single base

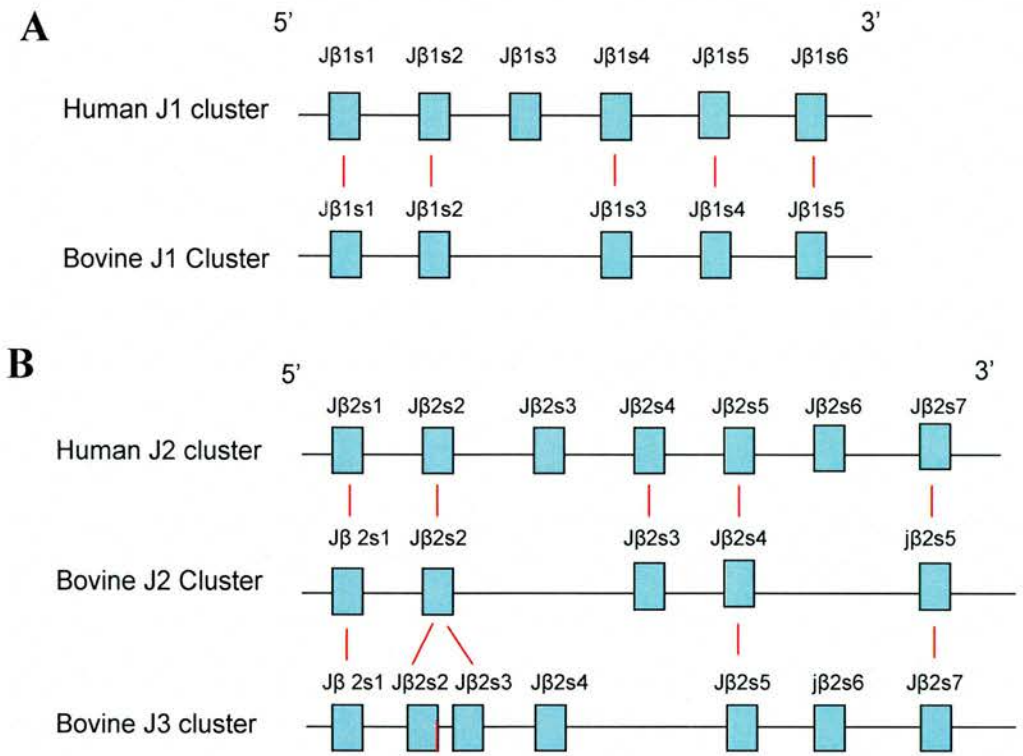


Figure 5.2 Comparative genomic organisation of the human and bovine Jβ gene clusters. Schematic representation of the order of Jβ genes in the aligned (A) human Jβ1 and bovine Jβ1 clusters and (B) human Jβ2 and bovine Jβ2 and Jβ3 clusters. Red lines indicate orthologous human and bovine Jβ genes (Appendix I and Table 5.4). The boxes representing genes, and the distances between the genes are not shown to scale.

pair deletion, which makes the genomic sequence non-functional. Jβ1s2 also lacked a consensual RSS nonamer, as did Jβ1s5 (Table 5.4), indicating that these germline sequences are ORF pseudogenes, unable to undergo somatic recombination. Assuming these are not due to sequencing errors in the genomic data, identification of sequences with identity to these two genes in rearranged TCRβ transcripts (J2 for Jβ1s2 and J9 for Jβ1s5) infers that these genes have allelic forms, of which at least one is functional. A precedent for this exists with human Jβ2s7, which has two alleles, one of which is functional and the other an ORF pseudogene (Folch and Lefranc, 2000a). Thus all of the Jβ genes present in the current bovine genome have been identified in TCRβ transcripts and vice versa.

Table 5.4 Bovine Jβ genes in the current bovine genome assembly (overleaf). The 17 Jβ genes have been ascribed names (column 1) according to convention (Kazatchkine, 1995). For each gene the corresponding temporary name given to matching sequences described from cDNA analysis is shown in column 2 and the orthologous human gene in column 3. The position of each gene in scaffold NW_931380.1/Bt4_WGA850_2. is detailed in column 4. The nucleotide sequence of the RSS and the Jβ gene and the translated amino acid sequence of the Jβ gene are displayed in columns 5 and 6 respectively. Genomic bovine Jβ1s2 and 1s5 are ORF pseudogenes as they have non-consensual nonamer RSS sequences (*), and Jβ1s2 has a frameshift due a single base pair deletion in the Jβ region. Based of the sequence of J2 derived from cDNA, a thymidine (shown in parentheses) has been introduced in Jβ1s2 (†).Conservation of heptamer sequence suggests that for several Jβ genes the RSS spacers contain 13 rather than 12 base pairs (‡).

Gene Name	Temp. Name	Human orthol-ogue	Location in contig NW_931380.1/Bt4-_WGA850_2	RSS Nonamer	Spacer	Heptamer	Jβ gene - nucleotide sequence (upper) and translated amino acid (lower)
Jβ1s1	J14	Jβ1s1	132,880-132,935	GCTTTCCTC	CTTGCTCCGCTT	CACTGTG	TG AAC ACT GAG GTT TTC TTT GGA AAA GGC ACC AGG CTC ACG GTT GTA G N T E V F F G K G T R L T V V
Jβ1s2	J2	Jβ1s2	132,731-132,801	CGTTGTAGG*	GTGGCTGTGTTC	TGATGTG	T TAT GAC TAT CAC TTC GGC CCA GGG ACC AAG CTG ACG GT(T)† GTA G Y D Y H F G P G T K L T V V
Jβ1s3	J11	Jβ1s4	131,854-131,932	AGTTTCCCT	ACTAGTCTTTAG	GGTTGTG	TG ACT AAT GAT AGA CTC TAT TTC GGC AAC GGG ACG AAG CTT TCC GTC TTG G T N D R L Y F G N G T K L S V L
Jβ1s4	J9	Jβ1s5	131,579-131,881	CTCTGGGAT*	TTGCCACACCCG	CATTGTG	CTG TGG AGC AAC CAG GCA CAG AAC TTT GGA CAC GGG ACC TGG CTC GCC GTC CTG G L W S N Q A Q N F G H G T W L A V L
Jβ1s5	J16	Jβ1s6	130,781-130,861	GGTTTTGCC	ACAGCCACCTGC	AGCTGTG	G GCC TAT AAT TCG CCC CTC CAC TTT GGA ATC GGC ACC AGG CTC ATC GTG ACA G A Y N S P L H F G I G T R L I V T
Jβ2s1	J7	Jβ2s1	123,057-123,134	GAATTCTTG	CTAGCCCCTTCC	CACTGTG	C TCC TAT GAG GAG CAG CAC TTC GGG CCA GGG ACC AGG CTC ATT GTG CTA G S Y E E Q H F G P G T R L I V L
Jβ2s2	J10	Jβ2s2	122,860-122939	AAGTTTGCG	CCTGGGTCCCCAG‡	GGCTGTG	TG AAC GCC GTG CAG CTG TAC TTT GGA GCT GGT TCC AAG CTG ACT GTG CTG G N A V Q L Y F G A G S K L T V L
Jβ2s3	J5	Jβ2s4	122,644-122,720	GGTTTTTGT	CCTGGGTCCCCG	GGCTGTG	AGC CAA AGC ACT CAG TAC TTC GGC GCG GGC ACC CGG CTG TCG GTG CTA G S Q S T Q Y F G A G T R L S V L
Jβ2s4	J8	Jβ2s5	122,527-122,602	GGTTTTTGC	GCGGGGCTGGGG	GGCCGTG	AC TCA GAG ACG CTG TAC TTC GGG CCG GGC ACG CGG CTC CAG GTG CTA G S E T L Y F G P G T R L Q V L
Jβ2s5	J13	Jβ2s7	122,242-122,317	GGGTTTGCT	TGCGGGACTGGGA‡	CTCTGTG	A GCC AGT GAG CGG TAT TTT GGC GCC GGC ACC AGG CTC ACG GTC ACA G A S E R Y F G A G T R L T V T

Table 5.4 Bovine Jβ genes in the current bovine genome assembly (1 of 2 pages). Notes on page 167

Gene Name	Temp. Name	Human orthol-ogue	Location in contig NW_931380.1/Bt4-_WGA850_2	5′flanking 12bp-RSS Nonamer	Spacer	Heptamer	Jβ gene - nucleotide sequence (upper) and translated amino acid (lower)
Jβ3s1	J15	Jβ2s1	113,562-113,839	GAATTCTTG	CTAGCCCCTTCC	CACTGTG	C TCC TAT GGG GAG CTG CAC TTC GGG CCA GGG ACC AGG CTC TCT GTG CTA G S Y G E L H F G P G T R L S V L
Jβ3s2	J10	Jβ2s2	113,366-113,445	AAGTTTGCG	CCTGGGTCCCCAG‡	GGCTGTG	TG AAC ACC CAG CCC CTG TAC TTT GGA GCT GGT TCC AAG CTG ACT GTG CTG G N T Q P L Y F G A G S K L T V L
Jβ3s3	J1	Jβ2s2	113,148-113,224	AGTTTTTGT	CCTGAGCCTCCG	GGCTGTG	AGC AAC AAC CCT CTG TAT TTT GGA GGG GGC ACT CGG CTG TTG GTA CTA G S N N P L Y F G G G T R L L V L
Jβ3s4	J4	Jβ2s3	113,001-113,077	TATTTTTGT	GCTGAGCCCTGA	GGCTGTG	AGC ACA GAC ACG CAG TAC TTC GGC GCG GGC ACC CGG CTG ACG GTG CTAG S T D T Q Y F G A G T R L T V L
Jβ3s5	J6	Jβ2s5	112,884-112,958	GGTTTTTGC	GCGAGGCTGGGG	GGCCGTG	ACT CAG ATC CAG TAC TTC GGG CCG GGC ACG CGG CTC CTG GTG CTA G T Q I Q Y F G P G T R L L V L
Jβ3s6	J12	Jβ2s6	112,782-112,862	GATTTTTGT	GGGGCGCCCCGA	GGCTGTG	C TTT GCA GGC GCC GCC CTG ACC TTC GGG GCC GGC AGC TGG CTG ACG GTG GTG G F A G A A L T F G A G S W L T V V
Jβ3s7	J3	Jβ2s7	112,556-112,631	GGGTTTGAG	TGCGGGGCTCTTC‡	TTCCGTG	T TCC TAT GAG CAG TAT TTC GGC CCA GGC ACC AAG CTC ACG GTC TTA G S Y E Q Y F G P G T K L T V L

Table 5.4 (continued) Bovine Jβ genes in the current bovine genome assembly (2 of 2 pages). Notes on page 167

5.3.3 Analysis of Dβ genes in the current bovine genome

Three sequences with significant homology to the available bovine germline Dβ gene sequences and/or the human Dβ genes were identified in the bovine genome. Due to the small size of Dβ genes and the potential for substantial modification of both the 3' and 5' ends during somatic recombination, no attempt had been made to identify bovine Dβ gene sequences during the analysis of TCRβ transcripts in Chapter 4. Each of the Dβ sequences identified was located 3' to one of the Jβ clusters on the scaffold NW_931380.1/Bt4_WGA850_2 and was named accordingly. Bovine Dβ1 shared 77% nucleotide identity with human Dβ1 and was of similar size (13bp in bovine compared to 12bp in human Dβ1). Bovine Dβ2 and Dβ3 shared 81% and 70% nucleotide sequence identity, respectively, with human Dβ2, and all three genes were 16bp in length. This level of nucleotide identity is comparable to that seen between human and porcine Dβ genes (Baron *et al.*, 2001). Therefore, it would appear that in the bovine there is one orthologous gene to human Dβ1 and two orthologous genes to human Dβ2, reflecting the situation of the Jβ gene clusters. The sequence details of the three Dβ genes are given in Table 5.5. Interestingly Dβ1 does not encode for a glycine (G) in the third reading frame; in other species studied all Dβ genes encode at least one glycine residue in each of the three reading frames (McCormack *et al.*, 1991).

5.3.4 Analysis of Cβ genes in the current bovine genome

Consistent with the report by Conrad *et al* (2002), 3 Cβ genes were identified in the bovine genome. Each of the Cβ genes was located 5' to a Jβ cluster on the scaffold NW_931380.1/Bt4_WGA850_2 (Appendix J and section 5.3.5), and was named accordingly. As in mice and humans, Cβ genes comprise four exons, but it was not possible to identify exon 2 for either Cβ2 or Cβ3, as the sequence in the regions predicted to contain these exons was undetermined. The aligned sequences of the three bovine Cβ genes are shown in Appendix J.

		5'flanking 12b-pRSS				3'flanking 23bp-RSS		
Dβ gene	Location on NW_931380.1 /Bt4_WGA850_2	Nonamer	Spacer	Heptamer	Dβ gene	Heptamer	Spacer	Nonamer
Dβ1	133,520-133,599	CGTTTTTGT	ATAAAGCTGTAA	CGTTGTG	GGG ACA GCA GGG C G T A G G Q Q G D S R	CACGGTG	ATTCAACCCTATGGGAATCCTTT	ACAAAAACC
Dβ2	123,710-123,792	CATTTTTGT	ATCACAGTATAA	CATTGTG	GGG ACT GGG GGG TGG G G T G G W G L G G G D W G V	CACAATG	ATTCAGTTAGAGGAAGTACTTTT	ACAAAAAGC
Dβ3	114,572-114,654	CATTTTTGT	ATCACAGTATAA	CATTGTG	GGA CTT CGG GGG GGG C G L R G G D F G G G T S G G	CACGATG	ATTCAGTTAGAGGAAGTGCTTTT	ACAAAAAGC

Table 5.5 Bovine Dβ genes identified in the current genome assembly. The location on the NW_931380.1/Bt4_WGA850_2 scaffold (column 2), nucleotide sequence of the 5'flanking 12-bp RSS (columns 3), nucleotide sequence of the Dβ gene and its translation in all 3 reading frames (column 4) and the nucleotide sequence of the 3'flanking 23-bpRSS are shown for each bovine Dβ gene identified in the current genome assembly.

The bovine Cβ1 and Cβ2 genes are very closely related, sharing >99% nucleotide sequence identity, whilst Cβ3 is more divergent with approximately 96% identity to the other two Cβ genes (Appendix J).The 3 bovine Cβ genes are more similar to each other than to either of the human Cβ genes (Appendix J) suggesting that there has been a process of intra-species homogenisation of the bovine Cβ gene sequences, such has been reported previously from humans and mice following comparison of their Cβ gene sequences (Glusman *et al.*, 2001).

5.3.5 Organisation of the bovine Dβ, Jβ and Cβ genes

The Dβ, Jβ and Cβ genes identified in the bovine genome are located in three, tandemly located DJC clusters on scaffold NW_931380.1/Bt4_WGA850_2. Although each cluster has the orthodox arrangement of several (five or seven) Jβ genes located between a 3' Dβ and a 5' Cβ gene, the presence of three instead of two DJC clusters is unique amongst the mammalian TCRB loci that have so far been characterised - human (Rowen *et al.*, 1996), murine (GenBank accession number AE000665), rat (Williams *et al.*, 1991) and porcine (GenBank accession number AB079894).

By describing human orthologues of the bovine Jβ (section 5.3.2) and Dβ (section 5.3.3) genes and comparing the nucleotide sequences of the bovine Cβ genes to each other (section 5.3.4) it would appear that the third DJC cluster has arisen from duplication of ancestral Cβ1, Dβ2 and Jβ2 cluster genes. The results of a dot-matrix analysis of scaffold NW_931380.1/Bt4_WGA850_2 versus itself with the TCRβ genes masked, is consistent with a duplication of the region incorporating Cβ1, Dβ2 and the Jβ2 cluster (Dr. J. Aerts, Roslin Institute, personal communication).

5.3.6 Regions of the bovine and human TCRB loci display conserved synteny.

The current map of bovine chromosome 4 places 6 scaffolds - NW_931374.1/Bt4_WGA844_2 to NW_931379.1/Bt4_WGA849_2 - none of which contain TCRβ genes, in between scaffolds NW_931373.1/Bt4_WGA843_2 and NW_931380.1/Bt4-_WGA850_2. These 6 scaffolds have a combined length in excess of 2,000Kb and the interposition of such a large insertion within the bovine TCRB locus would seem unlikely. One of these scaffolds, NW_931377.1/Bt4_WGA847_2, contains the muscle chloride channel 1 (CLCN1) gene, which is located outside of the TCRB locus in both humans and sheep, in positions that display conserved synteny (Antonacci *et al.*, 2001; Di Meo *et al.*, 2000). Additionally, as noted in Table 5.1, the TCRβ genes located on scaffolds NW_931372.1/Bt4_WGA842_2 and NW_931380.1/Bt4_WGA850_2 (except VB20a) appear to be in the reverse orientation whilst the TCRβ genes located on NW_931373.1/Bt4_WGA843_2 are in the forward orientation. In both the murine and human TCRB loci all of the TCRβ genes, with the exception of the orthologous human Vβ20 and murine Vβ14, have the same orientation (Rowen *et al.*, 1996, GenBank accession numbers AE000663-AE000665). These observations suggest that the order and orientation of scaffolds mapped to this area of chromosome 4 may require modification.

If the orientation of NW_931372.1/Bt4_WGA842_2 and NW_931380.1/Bt4-_WGA850_2 are inverted so that all 3 scaffolds holding constituents of the bovine TCRB locus are brought into the same orientation, the organisation of the genes on these scaffolds (and contig NW_935493.1/BtUn_WGA10890_2) shows marked conservation in synteny of TCRβ genes with corresponding regions of the human TCRB locus, as shown in Figure 5.3.

There is a striking conservation of the order of orthologous Vβ genes, including the location of Vβ20 3' to the DJC clusters and in the opposite transcriptional orientation to the other TCRβ genes (section 1.2.3.1) and although there is no human orthologue of VβXs1, the position of this gene at the extreme 5' end of the Vβ cluster is consistent with the position of its murine orthologue, Vβ2 (GenBank accession

number AE000663 - Figure 1.4). The large region (~100Kb) of undetermined sequence in NW_931372.1/Bt4_ WGA842_2 prevents any conserved synteny that may exist in this region from being demonstrated. As described in section 5.3.2 and Figure 5.2 there is also substantially conserved synteny between the order of the bovine and human DβJβCβ genes.

Synteny also extends to non-TCR genes in and adjacent to the TCRB locus. Trypsinogen (T) genes, a dopamine-β-hydroxylase-like (DBH-like) gene and an ephrin type-b receptor 6 precursor (EPHB6) gene are all located in positions on the bovine scaffolds that correspond to their positions in the human TCRB locus (Figure 5.3).

Figure 5.3 Schematic representation of the human TCRB locus and components of the bovine TCRB loci (overleaf). The human (upper) TCRB locus is displayed in three segments running from the 5'end (upper panel) to the 3'end (lower panel). Bovine genome scaffolds NW_931372.1/Bt4_WGA842_2, NW_931373.1/Bt4_WGA843_2, NW_931380.1/Bt4_WGA850_2 and contig NW_935493.1/BtUn_WGA10890_2 (Lopes *et al.*) are displayed against corresponding regions of the human TCRB locus. Scaffolds NW_931372.1/Bt4_WGA842_2, and NW_931380.1/Bt4_WGA850_2 have been inverted as discussed above. Genes within the loci are shown according to the legend overleaf. Vβ genes represented by two coloured boxes have allelic forms of different functional status. Red lines indicate orthologous human and murine TCRβ genes as identified by nucleotide sequence identity (sections 5.3.1 - 5.3.4). The arrows adjacent to human Vβ20s1 and bovine VB20a indicate the opposite transcriptional orientation of these genes compared to the other TCRβ genes in the TCRB loci. Non-TCR genes, dopamine-β-hydroxylase like (DBH-like), ephrin type-b receptor 6 precursor (EPHB6) and trypsinogen genes in and adjacent to the TCRB locus have been shown. The boxes representing genes, and the distances between the genes are not shown to scale. Exons are not shown. The format of the figure follows that of the IMGT database (http://imgt.cines.fr).

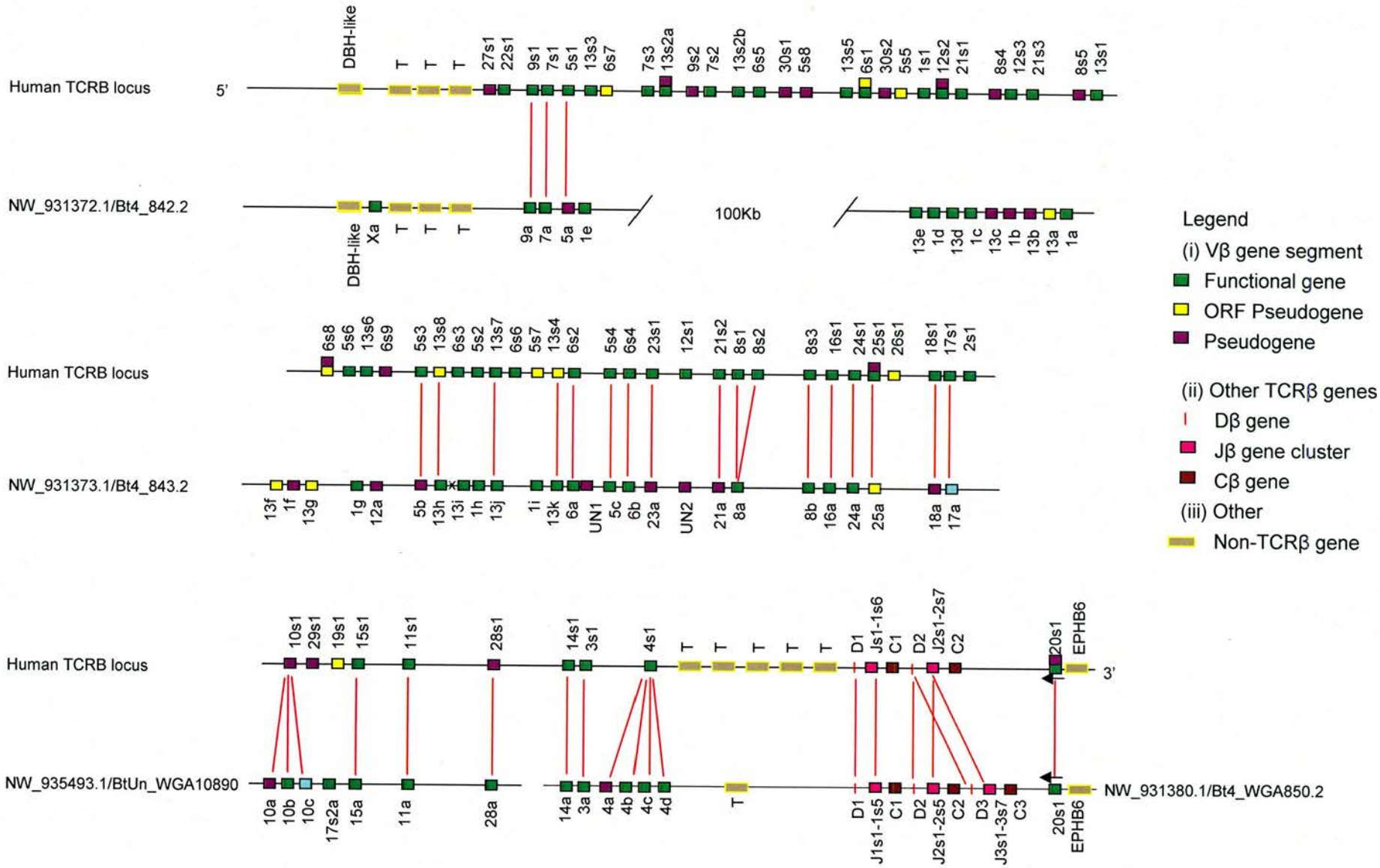


Figure 5.3 Schematic representation of the human TCRB locus and components of the bovine TCRB loci. Notes on previous page.

5.4 Discussion

In this chapter the current (second) bovine genome assembly has been examined for genomic sequence and organisation of the TCRB locus. At present the sequence of the TCRB locus is incomplete and fragmentary, with TCRβ gene sequences distributed over twenty-one contigs and five scaffolds; these results therefore represent a work 'in progress' and will require reviewing in the future as more genomic sequence becomes available. The imminent release of a third assembly (Dr Aerts, personal communication) is eagerly anticipated and will hopefully consolidate, if not complete, the genomic sequence of the TCRB locus and allow more definitive conclusions to be made.

The primary motivation for examining the genome assembly for TCRB locus sequence was to further characterize the repertoire of bovine TCRβ genes. Despite the incomplete status of the TCRB locus sequence, 123 Vβ sequences were identified, which represents a dramatic expansion compared to that of either the human or murine repertoires which are composed of 65 and 34 Vβ genes respectively (Bosc and Lefranc, 2000; Folch and Lefranc, 2000b). Much of this expansion is attributable to extensive duplication in certain subfamilies - most notably Vβ10 (16 genes), Vβ1 (29 genes) and Vβ13 (36 genes). Several novel bovine subfamilies were identified in the genome including single members of the Vβ11, Vβ21, Vβ23 subfamilies and 3 members of the Vβ18 subfamily. Bovine orthologues of all human Vβ subfamilies except Vβ19, 22, 26, 27, 29 and 30 were identified, as was an orthologue of murine Vβ2, which is absent in humans.

Accurate description of the repertoire of functional genes is essential for development of effective techniques to analyze TCR in functional studies. Acquisition of the full genomic sequence of most (114) of the Vβ genes identified enabled their functional status to be predicted by analysis of the DNA sequence. Of the genes identified 65 are predicted to be functional (56%), 35 are predicted to be pseudogenes (33%), 13 are predicted to be open-reading frame pseudogenes (11%) and 1 is an orphan pseudogene (1%). However, some caution must be exacted in

these predictions as (i) there may be functional allelic forms of some of the non-functional Vβ genes described in the genome, (ii) single base pair errors in sequencing are often sufficient to change the functional classification of Vβ genes and (iii) although generally accurate, prediction of functional status from analysis of DNA sequence is not always correct; there are two human Vβ genes (Vβ13s4 and Vβ6s8) which at the DNA level are predicted to be functional but protein modelling demonstrates are incapable of correct protein folding and are therefore non-functional (Rowen *et al.*, 1996).

The percentage of predicted Vβ pseudogenes present in the current bovine genome (44%) is higher but comparable to that in the murine (39%) or human (37% - including the 7 orphan pseudogenes on chromosome 9 - Robinson *et al.* (1993), Rowen *et al* (1996), GenBank accession number AF029308) Vβ germline repertoires. As in humans and mice, there are some bovine subfamilies, Vβ18, 21, 23 and 25, that at present do not contain any known functional Vβ genes. In humans these 4 families all have functional members, suggesting loss of function of the bovine orthologues has occurred subsequent to human/artiodactyl divergence 65-80 MYA (Li *et al.*, 1990). Although the Vβ12 genes identified in the bovine genome are also non-functional, analyses of cDNA have revealed two Vβ12 sequences with no obvious sequence defects.

In the human TCRB locus 22 Vβ 'relics' (i.e. genes with major lesions) were identified by application of a modified Smith-Waterman algorithm (Rowen *et al.*, 1996). The two severely mutated bovine Vβ genes named Untitled 1 and 2 probably represent 'relics' of bovine Vβ12 subfamily members, and by applying the modified Smith-Waterman algorithm in future work further 'relics' may be identified.

Approximately 35% (29 out of 84) of the Vβ genes identified from cDNA analysis had no matching genomic sequence in the current assembly. Genomic sequences matching 3 of these genes (Vβ2s1, 2s3 and 13s2) were subsequently identified in the WGS trace sequence archive. Conversely, approximately 19% (12 out 65) of the functional Vβ genes identified in the genome did not have matching cDNA

sequences. Thus, at present it would appear that results from cDNA analysis provide a more comprehensive coverage of the functional bovine TCRβ repertoire than the genomic data. Although this may simply reflect the current incomplete TCRB locus sequence in the genome, a potential confounding factor may be the presence of insertion-deletion related polymorphisms (IDRPs). TCRB locus IDRPs are known to exist in both humans and mice (Behlke *et al.*, 1986; Haqqi *et al.*, 1989; Jouvin-Marche *et al.*, 1989; Noonan *et al.*, 1986; Pullen *et al.*, 1990; Seboun *et al.*, 1989; Zhao *et al.*, 1994) and can result in deletion of up to 50% of the germline Vβ repertoire. It was notable that the majority (19 out of 26) of Vβ genes identified from cDNA analysis that lacked a matching genomic sequence, were members of the large multimember Vβ1, 2, 4, 10 and 13 subfamilies, as were 10 out of 11 of the genomic Vβ genes predicted to be functional that lacked a matching cDNA sequence. The discordance in the representation of the members of these large subfamilies identified from cDNA and genome analysis may indicate that the coverage of these subfamilies is still incomplete and their membership may be considerably larger than already appreciated.

Seventeen Jβ genes, 3 Dβ genes and 3 Cβ genes were identified in the genome. In contrast to the situation with the Vβ genes, all of the Jβ genes present in the genome had been identified in cDNA analysis and vice versa, suggesting comprehensive cDNA and genomic coverage of bovine Jβ has been achieved. Two of the genomic Jβ sequences (Jβ1s2 and Jβ1s5) were classified as ORF pseudogenes, but as they have been identified in cDNA it is assumed functional alleles of these genes must exist. All of the Dβ, Jβ and Cβ genes were present on a single scaffold - NW_931380.1/Bt4_WGA850_2 - which also contained a trypsinogen gene and the Vβ20 gene, which in humans and mice reside 5' and 3' to the DJC region. Thus, apart from a few minor areas of undetermined sequence within the scaffold, the entire DJC region of the bovine TCRB locus would appear to be present in the current assembly. The observed organisation of the bovine DJC genes into 3 clusters is unique amongst the mammals for which this region has been characterised; human - Rowen *et al* (1996), rat - Williams *et al* (1991), pig - (GenBank accession number AB079894) and mouse - (GenBank accession number AE000665).

However, southern blot analysis of genomic DNA from 2 rabbits identified the presence of 3 Cβ genes in one individual, which was corroborated by subsequent genomic sequencing (Komatsu *et al.*, 1987). The existence of a third bovine Cβ had been indicated by Conrad *et al* (2002) from preliminary analysis of a BAC clone containing TCRB locus genomic sequence. Analysis of TCRβ chain transcripts in the present study (Chapter 4) had also suggested the existence of two bovine Jβ clusters orthologous to the human Jβ2 cluster. The deposited genomic sequence (GenBank accession number AF453325) analysed by Conrad and colleagues matched part of the DJC region but did not contain the Dβ1 gene or the first four genes in the Jβ1 cluster; therefore the results presented here provide the first description of the entire bovine DβJβCβ region. The extra DJC cluster appears to have arisen by duplication of a region incorporating the Cβ1 gene, the Dβ2 gene and the Jβ2 cluster. Interestingly, this region has been deleted in New Zealand White mice (NZW), with an unequal cross-over event invoked as the most likely explanation (Kotzin *et al.*, 1985; Noonan *et al.*, 1986). The absence of a third DJC cluster in the pig (GenBank accession number AB079894) suggests that the duplication occurred after suidae/ruminant divergence 60-65 MYA (Kumar and Hedges, 1998). It will be interesting in the future to determine if the duplication is present in other ruminants such as the sheep and goat or if it is unique to cattle.

Forty-six of the Vβ sequences and all of the Dβ, Jβ and Cβ sequences were located on four scaffolds which have been mapped to chromosome 4. Location of the TCRB locus on chromosome 4 is consistent with results from fluorescent in-situ hybridisation (FISH) analysis (Antonacci *et al.*, 2001) and, as in humans, places the TCRB and TCRC loci on the same chromosome (Miccoli *et al.*, 2003). The location of one Vβ sequence (VB13l) on a scaffold mapped to chromosome 5 is presumably the result of an inter-chromosomal translocation event, as has occurred in humans (Robinson *et al.*, 1993; Rowen *et al.*, 1996). In contrast to humans, where multiple Vβ genes and a trypsinogen gene have been translocated from chromosome 7 to 9, possibly in multiple translocation events (Rowen *et al.*, 1996), only one Vβ gene has been identified on bovine chromosome 5 although 2Kb downstream there is a 19Kb region of undetermined sequence that may contain other Vβ elements.

The remaining 76 Vβ sequences identified were located on contigs that have yet to be mapped onto chromosomes. If it is assumed that all of these contigs will eventually map to the TCRB locus on chromosome 4, the locus would already be ≈650Kb in length, equivalent in size to both the human and murine TCRB loci (Rowen *et al.*, 1996, GenBank accession numbers AE000663-AE000665). The number of trypsinogen genes (8 in humans and 20 in mice) and Vβ genes present in the murine and human TCRB locus are inversely proportional (Glusman *et al.*, 2001). This inverse proportionality would appear to extend to the bovine TCRB locus in which only 4 trypsinogen genes have been identified; this is likely to represent the full complement of trypsinogen genes within the bovine TCRB locus, since sequence for the areas of the bovine TCRB locus corresponding to those in which trypsinogen genes are located in mouse and man appears to be complete.

Conservation in synteny of the TCRB locus and adjacent gene loci between humans and mouse is well established (Glusman *et al.*, 2001; Lai *et al.*, 1988). Previous studies using FISH have described conserved arrangement of the TCRB locus relative to adjacent gene loci between humans and ruminants (Antonacci *et al.*, 2001; Di Meo *et al.*, 2000). The availability of genomic sequence data has permitted analyses of synteny at a higher level of resolution and reveals a striking level of conservation in the organisation of genes within the TCRB locus between cattle and humans. Comparison of Figures 1.4 and 5.3 shows that the conserved synteny between bovine and human TCRB loci exceeds that displayed between the human and murine loci, consistent with the closer evolutionary relationship of cattle and humans (Kumar and Hedges, 1998; Li *et al.*, 1990). Direct interpretation of the conserved synteny between the human and bovine Vβ gene regions (as shown in Figure 5.3) would suggest that bovine Vβ19, 22, 26, 27 and 29 genes are absent from the germline repertoire. However, duplications of Vβ genes can distort apparent conserved synteny and such conclusions remain speculative in the absence of the complete TCRB locus sequence.

Extensive duplication of portions of the 5'end Vβ region and subsequent gene divergence or conversion has made reconstruction of a convincing evolutionary

history of the human TCRβ locus difficult (Glusman *et al.*, 2001; Rowen *et al.*, 1996). However, knowledge of the organisation and sequence of germline murine and human Vβ genes has been useful in analysing the evolutionary dynamics of Vβ subfamilies (Su and Nei, 2001). Although incomplete, the genomic sequence of the bovine TCRB locus provides insight into the evolutionary dynamics of bovine Vβ genes. A distinctive feature of the bovine Vβ genes is that the majority of the Vβ1 and Vβ13 genes are present in tandem (see Table 5.1 and Figure 5.3), suggesting that much of the expansion of these two subfamilies is a consequence of duplication of a Vβ1/Vβ13 cassette. Partial similarity of banding patterns in southern blots genomic bovine DNA hybridised with Vβ1 and Vβ13 specific probes had indicated that the members of these two subfamilies might be associated in the genome (Figure 4.10e). The bovine genome also contains tandem arrangement of Vβ1/Vβ12 and Vβ5/Vβ13 genes (Table 5.1). As discussed previously in Chapter 4, subfamilies Vβ1 and Vβ5 are very closely related as are the Vβ12 and 13 subfamilies, and segregation of the Vβ1/Vβ5 subfamilies and Vβ12/Vβ13 subfamilies may be an artificial consequence of the assigned subfamily membership of bovine Vβ genes being based on comparison to human Vβ sequence. Therefore, it is possible that tandemly arranged Vβ1/Vβ12 and Vβ5/Vβ13 in the bovine genome are descendants of duplications of the same cassette that have undergone subsequent divergence, more detailed phylogenetic analyses will be required to shed light on this issue.

Interestingly, in humans there are five repeats of a cassette containing Vβ5 and Vβ13 genes; Vβ1 and Vβ12s2 are also tandemly organised in the genome (Rowen *et al.*, 1996). It could therefore be conjectured that an ancestral Vβ1-5/Vβ12-13 cassette has been duplicated and subsequent evolution has resulted in the various tandem duplications evident in the cattle and human TCRB loci. Unlike the duplicated bovine cassette, some of the human Vβ5/Vβ13 duplications also contain a Vβ6 gene (Rowen *et al.*, 1996). These duplications have been estimated to have occurred 32-24MYA (Su and Nei, 2001), subsequent to primate/artiodactyl divergence (65-80 MYA - Kumar and Hedges, 1998; Li *et al.*, 1990). Thus duplication of the cassette in the two species must have occurred independently. The occurrence of parallel but

separate duplications in human and bovine Vβ repertoires raises fascinating questions about the evolutionary pressures that influenced these duplications.

Identification of numerous pairs of bovine Vβ genes (10) which share >98% nucleotide similarity (Appendix G) indicates that duplications of some bovine Vβ genes have occurred very recently. A large duplication in the murine TCRα, which occurred only 4-8 MYA has created 40 pairs of Vα genes which share ≈97% nucleotide identity (Glusman *et al.*, 2001). An alternative for some of these pairs of bovine Vβ genes may be homogenisation due to gene conversion events; evidence for which has been presented to explain similarity between human Vβ13s2a and 13s2b and has also been recognised in murine Vα and primate Vβ genes (Funkhouser *et al.*, 1997; Glusman *et al.*, 2001). The existence of distinct Vβ genes that share >98% nucleotide sequence identity complicates cDNA and genome sequence comparisons since allelic variants usually also show a similar low level of sequence difference.

Although the genomic sequence of the TCRB locus is incomplete, the results presented in this chapter reveal that (i) there is a massive germline Vβ gene repertoire in the bovine, (ii) there are three DJC clusters, as had been suggested previously (Conrad *et al.*, 2002) and (iii) there is marked conserved synteny in the organisation of the bovine and human TCRB loci.

6 Chapter 6 - Immunodominance, clonality and TCRβ chain expression in the bovine CD8$^+$ T-cell response to *Theileria parva*

6.1 Introduction

Theileria parva is a tick-transmitted parasite that causes a severe and usually fatal lymphoproliferative disease of cattle known as East Coast Fever (ECF) in eastern, central and southern Africa. ECF has significant economic impact on livestock agriculture in the affected region, causing annual losses estimated in 1989 to be approximately $168 million (Mukhebi *et al.*, 1992). These losses are due to a combination of the deaths of >1 million cattle/year (Norval *et al.*, 1992), reduced production and the cost of acaricides for tick control, which is currently the principal means of prevention. Due to issues of cost, availability, increasing resistance and environmental concerns, continued use of acaricides is considered to be unsustainable. An alternative control measure, immunisation by the infection and treatment (I & T) protocol (Radley *et al.*, 1975a), has had little impact because of the lack of local infrastructure for the production and distribution of the live sporozoites, which are required for immunisation. Therefore, current research efforts are directed towards the development of an effective subunit vaccine (Graham *et al.*, 2006; McKeever *et al.*, 1999; Musoke *et al.*, 2005).

Cattle that recover naturally from infection or are immunized by the I & T protocol are solidly immune to challenge with homologous and certain heterologous parasite strains (Radley *et al.*, 1975a). Kinetic and adoptive cell transfer studies have demonstrated that MHC class I-restricted CD8$^+$ CTL that target schizont-infected lymphocytes are prominent in mediating this immunity (McKeever *et al.*, 1994; Morrison *et al.*, 1987). Furthermore, there is a strong correlation between the specificity of the CTL response and cross-strain protection (Taracha *et al.*, 1995a). Thus CD8$^+$ T-cell antigens are obvious candidates for inclusion in a *T. parva* subunit vaccine.

The strain-specificity of the CTL response (Goddeeris *et al.*, 1986; Goddeeris *et al.*, 1990; Morrison *et al.*, 1987) together with evidence that there are hierarchies amongst both MHC-haplotypes (Morrison, 1996a; Morrison *et al.*, 1987) and epitopes (Taracha *et al.*, 1995b) in their ability to elicit $CD8^+$ T-cell responses indicate that *T. parva*-specific CTLs are focused on a limited number of immunodominant epitopes that exhibit polymorphism between different strains.

Immunodominance has been described in the $CD8^+$ T-cell responses against many viruses, bacteria and tumours (Yewdell and Bennink, 1999) and appears to be the result of numerous influences that operate concurrently (section 1.3). A thorough understanding of immunodominance is fundamental to the rational design of vaccines eliciting effective $CD8^+$ T-cell responses. In particular the inherent focusing of CTL against a limited number of epitopes, which is the defining characteristic of immunodominance, is a significant impediment to vaccine development where there is a need for broad CTL specificity to restrain emergence of 'escape mutants' (e.g. HIV) or to facilitate recognition of strain polymorphism in the immunodominant antigens (e.g. *T. parva*)

In several viral systems, including HCV and HIV, a diverse clonal composition and TCR repertoire within CTL responses against individual epitopes has been correlated with the ability to constrain 'escape mutants' that express epitope variants (Douek *et al.*, 2002; Lopes *et al.*, 2003; Meyer-Olson *et al.*, 2004; Turnbull *et al.*, 2006). Similarly, the clonality and heterogeneity in the TCR repertoire of CTL responses against defined *T. parva* epitopes may relate to their ability to recognise variant forms expressed by heterologous parasite strains and hence confer protection against these strains.

The recent identification of *T. parva* antigens recognised by CTL restricted by defined MHC haplotypes (Graham *et al.*, 2006) offers the opportunity to quantitatively assess immunodominance within the *T. parva* system. CTL-epitope mapping has identified two A10-restricted epitopes from the Tp2 antigen (Tp2.1 and Tp2.2) and a single A18-restricted epitope from the Tp1 antigen (Tp1.1). Using the

tools developed in Chapter 3, it is now also possible to analyze the diversity of the expressed TCRβ chains and clonal composition of CTL induced by these defined epitopes.

In this chapter is presented the analysis of immunodominance of $CD8^+$ T-cell responses to *T.parva* (Muguga) in animals homozygous for the A10 and A18 MHC class I haplotypes. The results show that the Tp2.2 epitope is highly dominant and the Tp2.1 epitope subdominant in the A10-restricted *T. parva*-specific $CD8^+$ T-cell response, whilst the Tp1.1 epitope is highly dominant in the A18-restricted response. The $CD8^+$ T-cell responses to both Tp2.2. and Tp1.1 appear to be polyclonal and express a diverse repertoire of TCRβ chains, but are dominated by a small number of large clonal expansions.

6.2 Materials and methods

6.2.1 Animals, immunisation and challenge

Animals 468 ($A18^+$ homozygous), 592 ($A10^+$ homozygous), 641 ($A18^+$ homozygous) and 1011 ($A10^+$ homozygous) were immunized using the CTVM *Theileria parva* (Muguga) sporozoite stabilate 71 and challenged as described in section 2.2.

6.2.2 Generation of $CD8^+$ T-cell enriched cell-lines and clones

$CD8^+$ T-cell enriched cell-lines were generated from PBMC by stimulation *in vitro* with irradiated autologous parasitized cells as described in section 2.3.3. Cell-lines were generated from animals 641 and 1011 prior to immunisation and from animals 468, 641 and 1011 >2 months after immunisation. The cell-line derived from 592 was established with PBMC collected >2 months after a second homologous *in vivo* challenge following the initial immunisation. By collecting PBMC at least 2 months after immunization or challenge, it was hoped that the resultant cell-lines would be representative of the $CD8^+$ T-cell memory pool. Approximately 90 clones were derived from each of the 4 post-immunisation cell-lines by limiting dilution (section 2.3.4).

6.2.3 Immunofluorescence staining and flow cytometry

The constituent cell populations of cell-lines obtained pre- and post-immunisation from animals 641 and 1011 were assessed as described in section 2.3.7 using the monoclonal antibodies (mAb) listed in Table 6.1. The phenotypes of clones that did not show specific cytotoxic activity for peptide-loaded MHC-matched Ta targets were assessed using monoclonal antibodies specific for CD3, CD4, CD8, WC1 (a gamma/delta T cell marker) and NKp46 (a marker for NK cells) . Any clones with a phenotype inconsistent with that of a $CD8^+$ T-cell were excluded from analysis. Clones that demonstrated specific cytotoxic activity against the appropriate peptide loaded MHC-matched targets were assumed to be $CD8^+$ T-cells.

6.2.4 Cytotoxicity assays

Cytotoxicity assays were performed as described in section 2.3.8. All cell-lines/clones were tested against autologous TpM. Cell-lines/clones from $A18^+$ animals were also tested against (i) 592 TpM (MHC-mismatched ($A10^+$) TpM), (ii) 468 Ta (MHC-matched ($A18^+$) Ta) and (iii) Tp1.1-loaded 468 Ta. Cell-lines/clones from $A10^+$ animals were also tested against (i) 641 TpM (MHC-mismatched ($A18^+$) TpM) and (ii) 592 Ta (MHC-matched ($A10^+$) Ta) peptide-loaded with either Tp2.1 or Tp2.2. Cytotoxic activity of 592 clones against the MHC-mismatched TpM and 468 clones against MHC-matched Ta was not examined.

6.2.5 Sequencing of TCRβ chains

cDNA from clones was generated as described in sections 2.4.1-2.4.3. The TCRβ chains were amplified using the Vβ subfamily-specific semi-nested PCR and/or the 'Pan-Vβ' semi-nested PCR protocols (section 2.4.4.2). The PCR products were sequenced (sections 2.4.9.1-2.4.9.2) and the TCRβ transcript sequences analysed as described in sections 2.5 and 4.2.2. The sequences of TCRβ chains have been presented in the format of Chothia *et al.* (1988) - section 2.5.4.

6.2.6 TCRβ chain CDR3 (CDR3β-HDA) heteroduplex assays

CDR3β-HDA was performed on cDNA from T-cell clones, cell-lines and PBMC collected from animal 592 immediately prior to a third *in vivo* challenge (pre-challenge) and at the peak of the response to this challenge (post-challenge) on day 10 (Morrison *et al.*, 1987) (McKeever *et al.*, 1994) as described in section 2.4.8.2. The PBMC were enriched for $CD8^+$ T-cells by lysis of $CD4^+$ γδ T-cells with antibody-directed complemented mediated lysis (section 2.3.2); the efficacy of this depletion was assessed by FACS analysis with mAb specific for CD4, WC1, CD8 and CD3 (section 2.3.7) prior to processing for CDR3β-HDA.

6.3 Results

6.3.1 $CD8^+$T-cell lines obtained post-immunisation but not those obtained pre-immunization are cytotoxic for targets presenting defined *T. parva* (Muguga)-epitopes

PBMC taken from animals 641 (homozygous for the A18 MHCI haplotype) and 1011 (homozygous for the A10 MHCI phenotype) pre- and post-immunisation with *Theileria parva* (Muguga) were stimulated *in vitro* with autologous *T. parva* (Muguga)-infected cells (TpM) and depleted of $CD4^+$ and γδ T-cells. The resulting $CD8^+$ enriched cell-lines were assessed for cytotoxic activity against (i) autologous TpM, (ii) MHC-mismatched TpM, (iii) MHC-matched *Theileria annulata*-infected cells (Ta) and (iv) MHC-matched Ta that had been pulsed with synthetic peptides representing the A10-restricted epitopes Tp2.1 or Tp2.2 or the A18-restricted epitope Tp1.1.

The cell-line obtained from animal 1011 ($A10^+$) prior to immunisation demonstrated cytotoxicity approaching 50% against the autologous TpM at effector:target ratios of 40:1 and 20:1 but no significant cytotoxic activity against a MHC-mismatched TpM (Figure 6.1a). In contrast, the cell-line obtained from the other naïve animal, 641 ($A18^+$), displayed equivalent levels of killing (approximately 30%) against both the autologous and MHC-mismatched TpM (Figure 6.1b). Neither of the pre-immunization cell-lines displayed any significant cytotoxic activity against the MHC-matched Ta targets, either in their native state or when loaded with peptide (Figure 6.1a and b).

The cell-lines derived from 641 and 1011 post-immunisation displayed maximum levels of cytotoxicity on the autologous TpM targets (~30% and ~50% respectively) similar to those detected prior to immunisation (although the 1011 cell-line gave higher levels of cytotoxicity at lower effecter:target ratios), but they did not kill MHC-mismatched TpM or MHC-matched Ta lines (Figure 6.1). In contrast to the pre-immunisation cell-lines, the post-immunisation cell lines from both animals demonstrated high levels of cytotoxic activity against the MHC-matched Ta targets

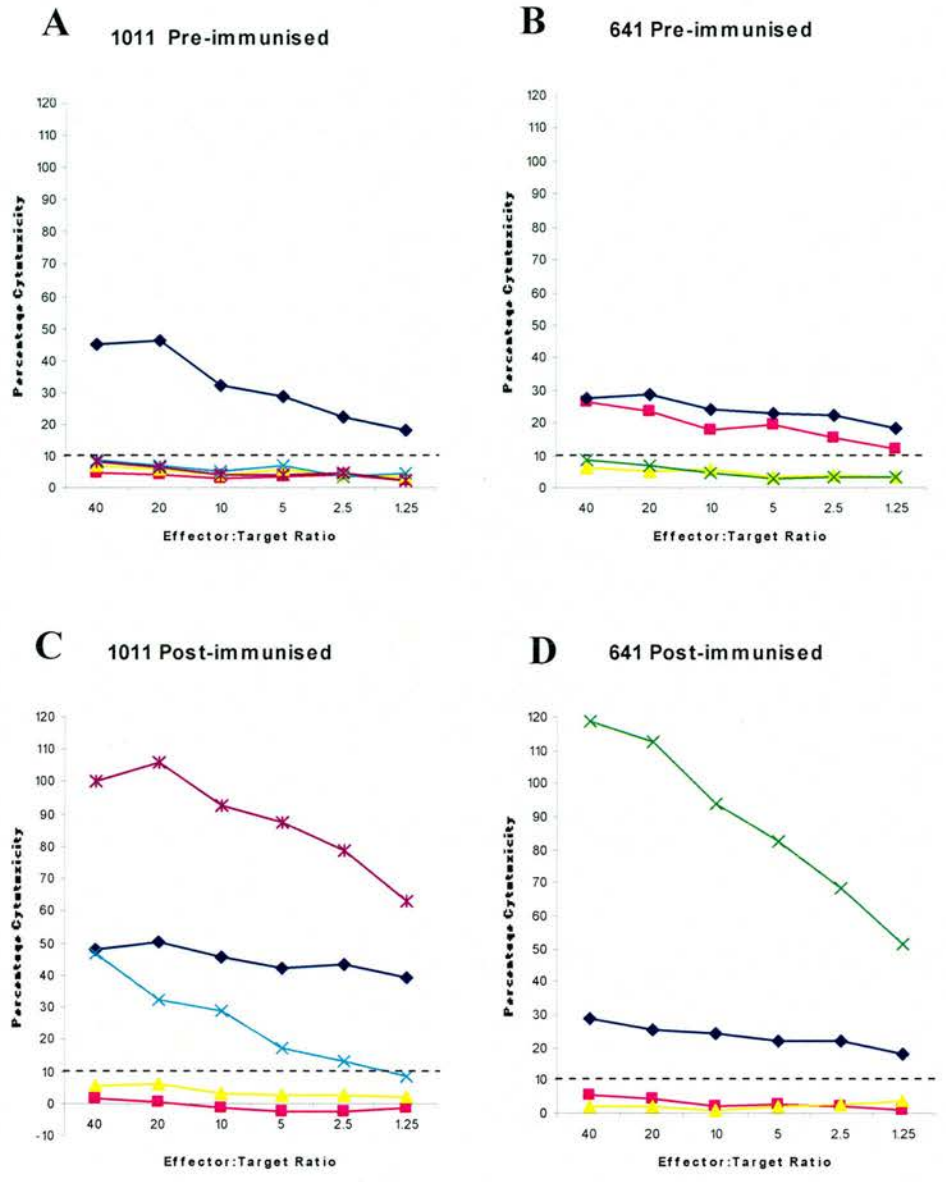


Figure 6.1 Cytotoxic activity of *in vitro* CD8+ T-cell enriched cell-lines derived from PBMC of (a) 1011 pre-immunisation, (b) 641 pre-immunisation, (c) 1011 post-immunisation, (d) 641 post-immunisation against: autologous TpM (dark blue line with diamonds) , MHC-mismatched TpM (pink line with squares), MHC-matched Ta (yellow line with diamonds), MHC-matched Ta loaded with Tp2.1 (light blue line with crosses), MHC-matched Ta loaded with Tp2.2 (black line with crosses), MHC-matched Ta loaded with Tp1.1 (green line with crosses). The level of cytotoxicity required to be considered significant (≥10%) is represented by the dashed line.

loaded with peptide epitopes. The post-immunisation cell line from the $A10^+$ animal (1011) gave maximum levels of killing of 47% and 106% of targets incubated with the Tp2.1 and Tp2.2 epitopes respectively (Figure 6.1c), while the line from the $A18^+$ animal (641) gave a maximal level of killing of 119% of the Tp1.1-loaded target (Figure 6.1d).

Thus, although both the pre- and post-immunisation cell-lines demonstrated cytotoxic activity against the MHC-matched TpM targets, only the post-immunisation cell-lines exhibited cytotoxicity against targets presenting the defined *T. parva* (Muguga) epitopes. The failure of the post-immunisation cell-lines to lyse the native MHC-matched Ta indicates the cytotoxicity against the peptide loaded targets was epitope-specific.

6.3.2 Phenotype pre- and post-immunisation cell-lines

The phenotype of the cell-lines derived from the pre- and post-immunisation PBMC were examined using a panel of mono-clonal antibodies in single colour FACS analysis (Table 6.1). Minor populations of B-cells and monocytes ($MyD\text{-}1^+$ cells) were evident in all four cell-lines, as were remnant populations of $CD4^+$ T-cells and $WC1^+ \gamma\delta$ T-cells following depletion of these populations by antibody-directed complement mediated lysis. The proportion of the populations that were $CD8^+$ varied between the cell lines, from 26.9% in the pre-immunisation 1011 cell-line to 84.6% in the post-immunisation 1011 cell-line.

However, the most notable difference between the pre- and post-immunisation cell-lines was the disparity in the number of T-cells ($CD3^+$) and NK-like cells ($NKp46^+$). Cell-lines obtained pre-immunisation were dominated by $NKp46^+$ cells (76.2% in 1011 and 64.6% in 641), with $CD3^+$ cells representing <30% of the population. In contrast, in both post-immunisation cell-lines $CD3^+$ cells constituted >80% of the cell population whilst $NKp46^+$ cell populations were comparatively minor (13.8% in 1011 and 6% in 641).

Antibody	Specificity and cell distribution	1011		641	
		Pre-immunisation	Post-immunisation	Pre-immunisation	Post-immunisation
IL-A12	CD4 - $CD4^+$ T-cells	9.2%	12.2%	7.2%	5.8%
CC15	WC1 - $WC1^+$ γδT-cells	8.8%	2.2%	3.1%	7.9%
IL-A51	CD8 - $CD8^+$ T-cells and NK-T cells	26.9%	84.6%	59.9%	69.8%
MM1A	CD3 - T-cells	30.5%	83.9%	29.0%	86.5%
NKp46-AKS1[i]	NKp46 - NK-like cells	76.2%	13.8%	64.6%	6%
IL-A24	MyD-1 - monocytes/macrophages/dendritic cells	<1%	ND	<1%	<1%
IL-A30	IgM - B-cells	2.8%	ND	2.2%	6.5%

Table 6.1 Phenotype of the *in vitro* cell-lines generated from pre- and post-immunisation PBMC from animals 1011 and 641.

6.3.3 Cytotoxic activity of *Theileria parva*-specific $CD8^+$ T-cell clones

Approximately 90 clones were generated from bulk $CD8^+$ T-cell-enriched cell-lines established from two immunised A10-homozygous (1011 and 592) and two immunised A18-homozygous animals (468 and 641), and their cytotoxic activity analysed. The results from this analysis are detailed in Appendix K.

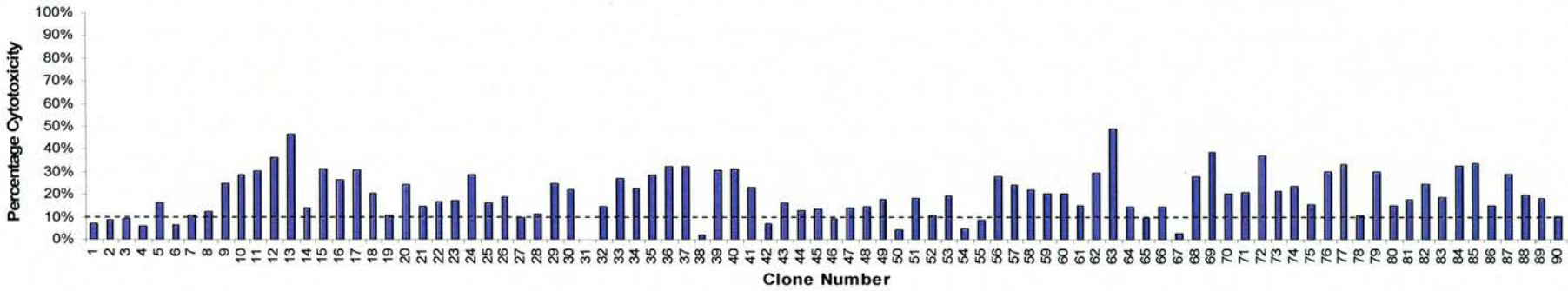
6.3.3.1 Tp2.2 is the immunodominant $CD8^+$ T-cell epitope in $A10^+$ animals immunised against *Theileria parva* (Muguga).

Over 80% of the clones generated from the two $A10^+$ animals (76/89 for 592 and 73/89 for 1011) demonstrated significant cytotoxicity against the autologous TpM targets (Figures 6.2a and 6.3a), with levels of killing ranging from 10.0% to 49.7% for 592 (Figure 6.2a) and 10.2%-79.0% for 1011 (Figure 6.3a). None of the 1011 clones exhibited significant cytotoxicity when tested against an MHC-mismatched ($A18^+$) TpM target (Figure 6.3a).

The majority of the clones from both 592 (71/89 - 80%) and 1011 (65/89 - 73%) demonstrated significant cytotoxicity against the Tp2.2-loaded $A10^+$ Ta target, whereas only 9/89 (10%) of clones from 592 and 2/89 (2%) of clones from 1011 exhibited significant cytotoxicity against the Tp2.1-loaded $A10^+$ Ta target (Figures 6.2b and 6.3b). As with the levels of cytotoxicity against the autologous TpM, there was considerable variation in the levels of cytotoxicity displayed against the peptide loaded Ta targets - e.g. amongst the 592 clones cytotoxic for the Tp2.2-loaded target percentage lysis varied from 63.7% (592.30) to 121.4% (592.73). With the exception of one clone from animal 592 (592.15), none of the clones exhibited significant cytotoxicity against both Tp2.1 and Tp2.2 (Figure 6.2b), confirming that the cytotoxic activity was epitope-specific (see section 6.3.4).

These findings indicate that Tp2.2 is highly dominant in the response of immunised $A10^+$ animals against *Theileria parva* (Muguga) as measured *in vitro*, whilst Tp2.1 is a subdominant epitope. The identification of clones not cytotoxic for either Tp2.1 or Tp2.2 presenting targets, some of which exhibited cytotoxicity against the autologous TpM target (e.g. 592.63), indicates the potential presence of additional subdominant epitopes in the response.

A



B

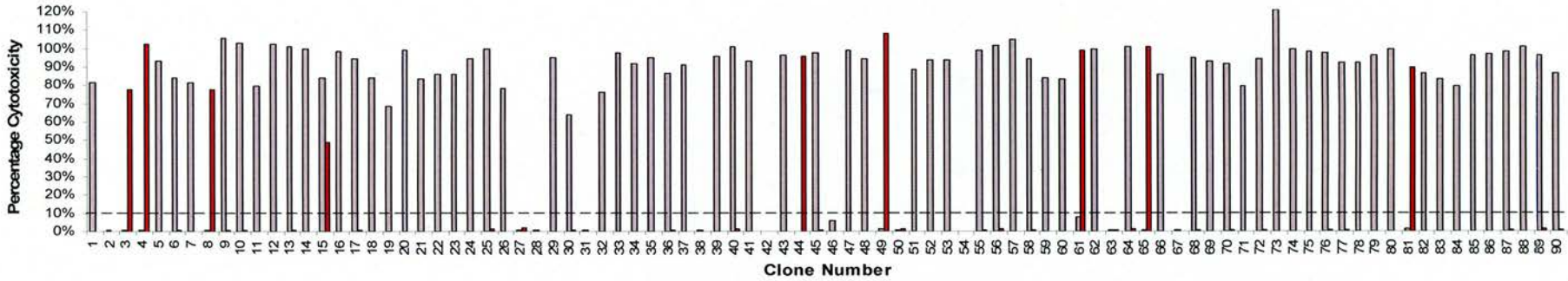
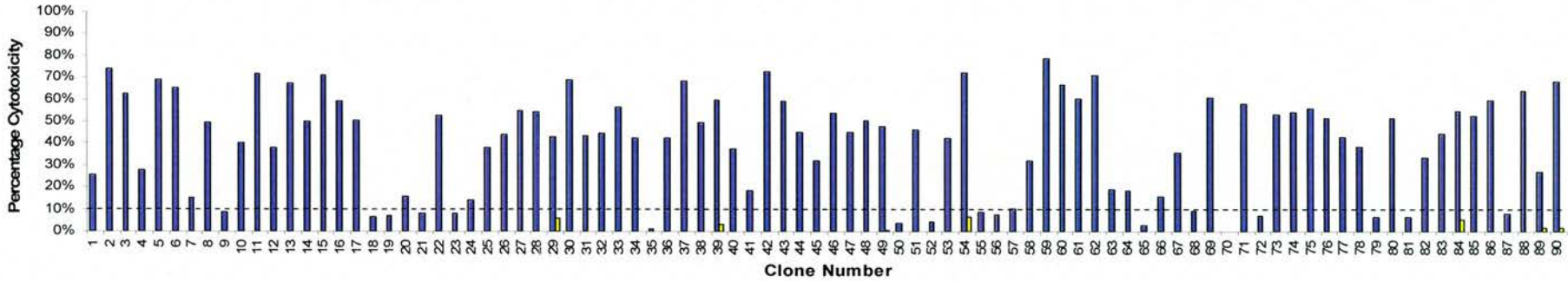


Figure 6.2 Cytotoxic activity of 592 clones. (a) The percentage cytotoxicity against the MHC-matched TpM target (blue bars).(b) The percentage cytotoxicity against the $A10^+$ Ta target peptide loaded with epitope Tp2.1 (red bars) and Tp2.2 (grey bars). The level of cytotoxicity required to be considered significant (≥10%) is represented by the dashed line. Cytotoxicity values <0% have been presented as 0% - actual values are given in Appendix K. No data is presented for clone 592.31.

A



B

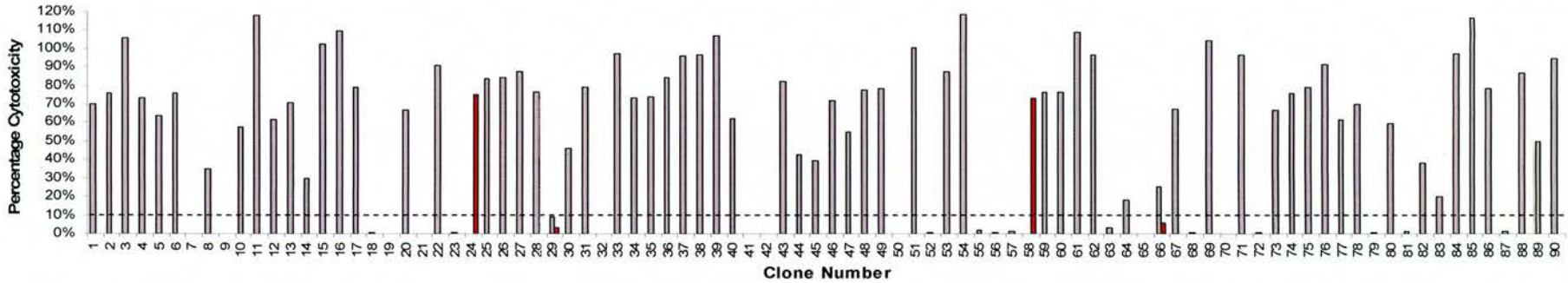


Figure 6.3 Cytotoxic activity of 1011 clones. (a) The percentage cytotoxicity against the MHC-matched TpM target (blue bars) and MHC-mismatched TpM target (yellow bars). (b) The percentage cytotoxicity against the $A10^+$ Ta target peptide loaded with epitope Tp2.1 (red bars) and Tp2.2 (grey bars). The level of cytotoxicity required to be considered significant (≥10%) is represented by the dashed line. Cytotoxicity values <0% have been presented as 0% - actual values are given in Appendix K. No data is presented for clone 1011.70.

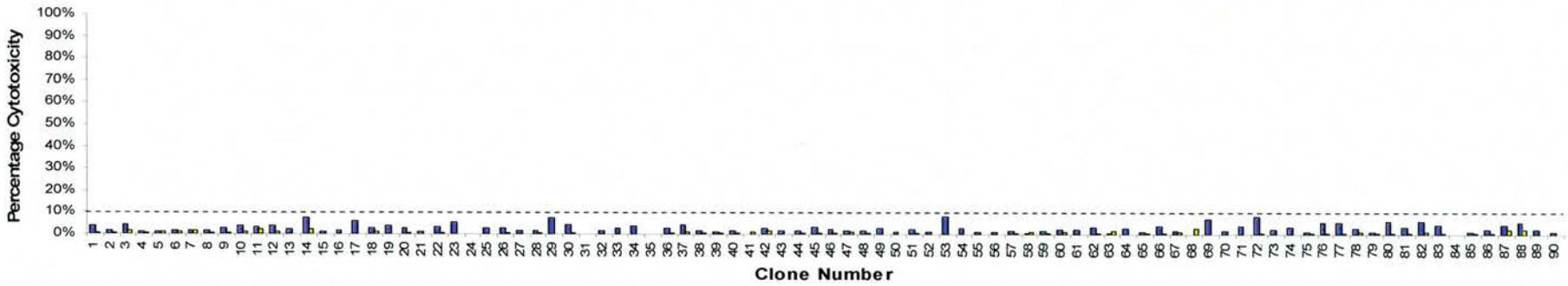
6.3.3.2 Tp1.1 is the immunodominant CD8$^+$ T-cell epitope A18$^+$ animals immunised against *Theileria parva* (Muguga).

The level of cytotoxicity displayed by the clones derived from the two A18-homozygous animals against the autologous TpM targets was generally low. In fact, none of the 468 clones achieved the 10% lysis considered to indicate significant cytotoxic activity and only 12 clones showed >5% cytotoxicity (Figure 6.4a). A minority (39/87 - 44.8%) of the 641 clones demonstrated significant cytotoxic activity against the autologous TpM target, but the maximum lysis achieved was only 33.3% (Figure 6.5a). None of the clones from either animal displayed any significant cytotoxic activity against the MHC-mismatched TpM target (Figures 6.4a and 6.5a).

Despite the low levels of killing detected on TpM targets, the majority of the clones from both 468 (72/90 - 80%) and 641 (64/87 - 74%) gave significant killing when assayed against the A18$^+$ Ta target loaded with the Tp1.1 epitope (Figures 6.4b and 6.5b). Again, the levels of killing varied markedly between clones - ranging from 13.9 to 125.2% for 468 clones and from 11.2 to 117.2% for 641 clones. None of the 641 clones showed any significant cytotoxic activity against the A18$^+$ Ta target in the absence of Tp1.1 peptide (Figure 6.5b), confirming their epitope specificity.

Thus, the Tp1.1 epitope was highly dominant in the response of both A18$^+$ *Theileria parva* (Muguga)-immunised animals.

A



B

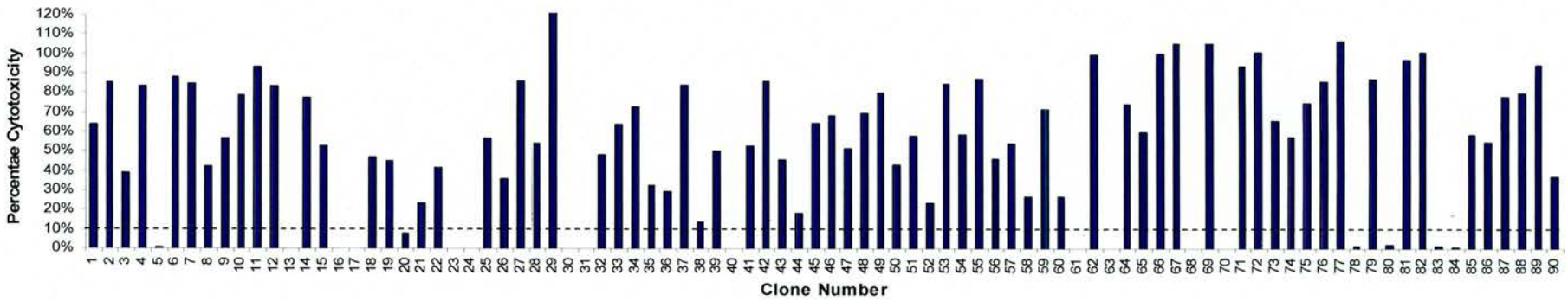
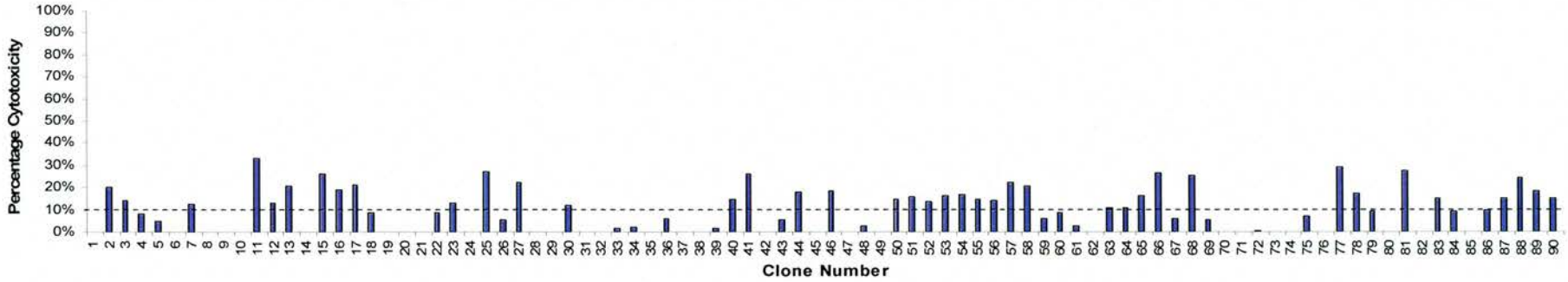


Figure 6.4 Cytotoxic activity of 468 clones. (a) The percentage cytotoxicity against the MHC-matched TpM target (blue bars) and MHC-mismatched target (yellow bars). (b) The percentage cytotoxicity against the $A18^{+}$ Ta target peptide loaded with epitope Tp1 (dark blue bars). The level of cytotoxicity required to be considered significant (≥10%) is represented by the dashed line. Cytotoxicity values <0% have been presented as 0% - actual values are given in Appendix K.

A



B

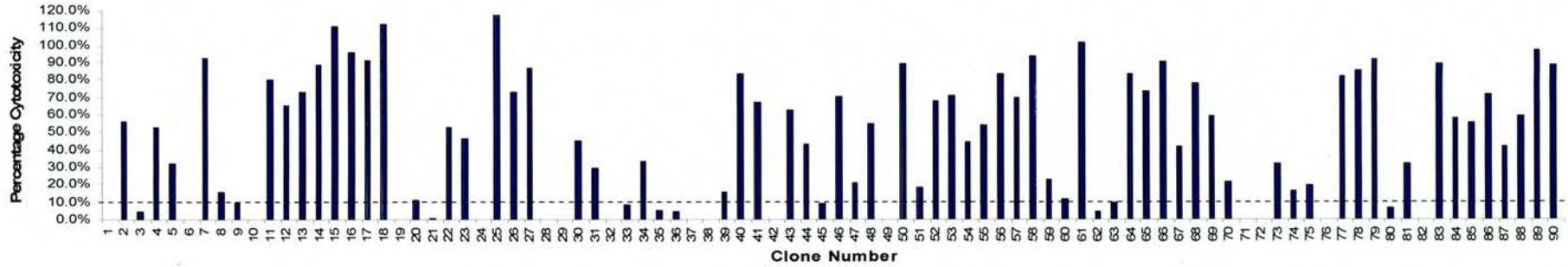


Figure 6.5 Cytotoxic activity of 641 clones. (a) The percentage cytotoxicity against the MHC-matched TpM target (blue bars) and MHC-mismatched TpM (yellow bars). (b) The percentage cytotoxicity against the $A18^+$ Ta target peptide loaded with epitope Tp1 (dark blue bars) and the unloaded $A18^+$ Ta target (purple bars). The level of cytotoxicity required to be considered significant (≥10%) is represented by the dashed line. Cytotoxicity values <0% have been presented as 0% - actual values are given in Appendix K. No data is presented for clones 641.19, 28 and 49.

6.3.4 Examination of TCRβ chain expression by clones

To determine the clonal composition of the *T. parva*-specific $CD8^+$ T-cell response represented in the *in vitro* cell-lines from these 4 animals, and the repertoire of TCRβ chains expressed by the responding clonotypes, the TCRβ chains expressed by the T-cell clones were sequenced following PCR amplification using the Vβ subfamily-specific primers and/or the 'PanVβ' primer (Appendix L). Sequences for a single predicted functional TCRβ chain was obtained for 310 of the 355 clones (87%) (Table 6.2).

Animal	468	592	641	1011	Total
Number of clones	90	89	87	89	355
Single functional TCRβ sequenced	67	78	84	81	310
No TCRβ chain sequenced	16	3	2	4	25
Dual functional TCRβ sequenced	7	8	1	4	20

Table 6.2 Summary of the results of sequencing the TCRβ chain expression by clones derived from animals 468, 592, 641 and 1011.

Some of these clones expressed additional TCRβ chains that were predicted to be non-functional (Appendix L - sequence in red script) due either to frameshifts in the CDR3 region or the expression of non-functional Vβ genes (Vβ10s6 and Vβ13s6 - section 4.3.1). For 25 clones (7%) no TCRβ chain sequence was obtained, whilst for the remaining 20 clones (6%) two predicted functional TCRβ chains were identified. In the majority of the latter clones, at least one of the TCRβ chains expressed was identical to a sequence identified among the clones expressing a single functional TCRβ chain (e.g. 592.41 expresses a $V\beta1^+$ chain identical to that in 592.75 and a $V\beta14^+$ chain identical to that in 592.26 and other clones), indicating that the expression of dual functional TCRβ chains by these 'clones' is predominantly attributable to their non-clonal nature. Analysis of TCRβ chain expression by clone 592.15 reveals that it expresses two predicted functional chains - a $V\beta2^+$ chain that is also expressed by 6 other clones that all display Tp2.2-specific cytotoxic activity, and a $V\beta14^+$ chain that is expressed by 8 other clones which all exhibit Tp2.1-specific cytotoxicity. The bi-clonality of this clone, indicated by the dual TCRβ

expression, provides an explanation for its ability to recognise both Tp2.1 and Tp2.2 epitopes (section 6.3.3.1). Only sequences of TCRβ chains expressed by clones with a single functional TCRβ chain were included in subsequent analyses.

6.3.5 Clonal composition of the response to immunodominant epitopes

The cytotoxicity results described in section 6.3.3 indicate that Tp1.1 and Tp2.2 are the immunodominant epitopes in the $CD8^+$ T-cell responses to *Theileria parva* (Muguga) in $A18^+$ and $A10^+$ animals respectively. Acquisition of the single predicted functional TCRβ chain sequence of 61 and 60 Tp2.2-specific CTL clones from animals 592 and 1011 respectively and 55 and 64 Tp1.1-specific CTL clones from animals 468 and 641 respectively, allowed characterisation of the clonality of the responses as detected in these *in vitro* cell-lines.

For each of the animals, the clones were organised into clonotypes on the basis of expression of TCRβ chains with identical sequence (Tables 6.3 and 6.4). The epitope-specific clones from animals 468, 592, 641 and 1011 belonged to 12, 17, 12 and 10 clonotypes respectively. As can be seen in Figure 6.4, where the cumulative percentage of clones is shown in relation to the number of clonotypes, for each animal a majority of the clones belong to a few large clonotypes. This is most pronounced in 641 and 1011, where >75% of the immunodominant epitope-specific CTL clones belong to only 2 clonotypes. The trend is less pronounced for animal 592, where the accumulation of the 8 most abundant clonotypes constituted 75% of the clones. The long 'tail' of the curves, particularly seen for 1011, 468 and 641, illustrate the presence of numerous less abundant clonotypes. Therefore, the CTL responses measured *in vitro* against the immunodominant epitopes appear to be polyclonal but dominated by a small number of large clonal expansions.

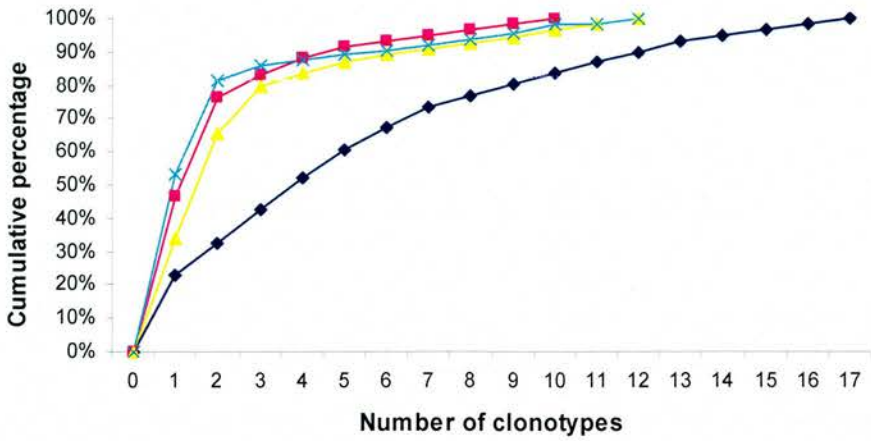


Figure 6.6 The response to the immunodominant A18-restricted (Tp1.1) and A10-restricted (Tp2.2) epitopes measured *in vitro* are dominated by a limited number of clonotypes. The cumulative percentages of Tp1.1 specific-clones from 641 (light blue), 468 (yellow) and Tp2.2-specific clones from 1011 (pink) and 592 (dark blue) are plotted against the number of clonotypes represented. Clonotypes are arranged in descending order of frequency.

6.3.6 TCRβ chain repertoires of Tp2.2 and Tp1.1-specific CTL clones

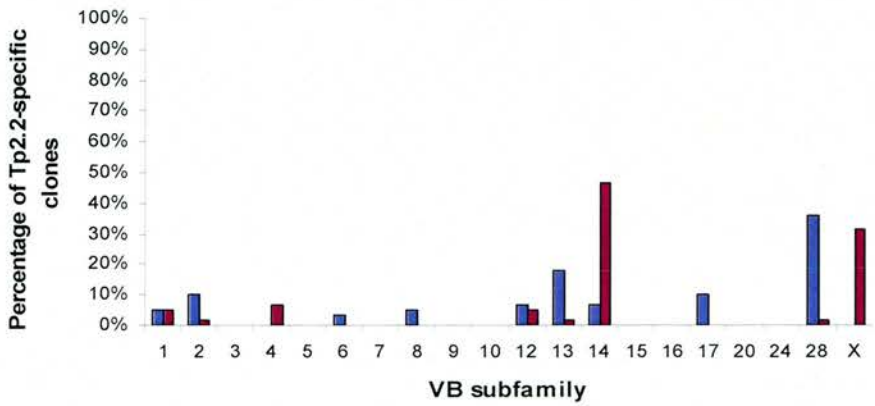
6.3.6.1 Diverse Vβ subfamily gene usage by Tp1.1- and Tp2.2-specific CTL clones

Vβ genes from 11 subfamilies were used by the 27 different Tp2.2-specific CTL clonotypes identified from 592 and 1011. As shown in Figure 6.7a, the profile of the Vβ subfamily usage by clones from these two animals was markedly different. Genes from Vβ subfamilies 6, 7 and 17 were expressed by clones from 592 but not 1011 and conversely Vβ4 and X were used by 1011 but not 592 clones. Furthermore, for Vβ subfamilies utilized by clones from both animals the frequency of expression was often dissimilar (e.g. Vβ14 and 28).

Similarly, the 24 Tp1.1-specific CTL clonotypes identified from 468 and 641 utilized Vβ genes from 11 subfamilies. Again, there was obvious discordance of Vβ

subfamily gene usage between clones from the 2 animals (Figure 6.7b), with Vβ2, 6, 20 and 24 genes used only by clones from 468 and Vβ4 and 16 gene usage being unique to clones from 641. Interestingly, due to the presence of single large clonotypes in both 468 and 641 the single member Vβ3 subfamily was used at high frequency by both 468 (34.5%) and 641 (54.7%) clones.

A



B

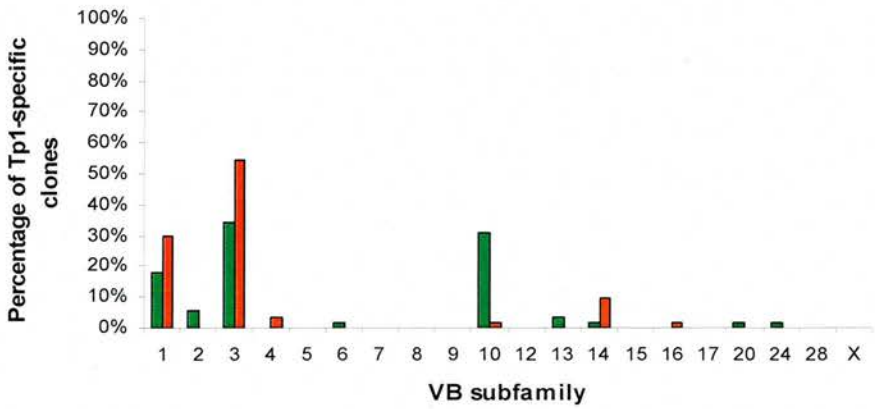


Figure 6.7 Percentage of immunodominant epitope-specific CTL clones expressing different Vβ subfamily genes. (a) Tp2.2-specific clones from 592 (light blue bars) and 1011 (purple). (b) Tp1-specific clones from 468 (Zhang *et al.*) and 641 (orange).

6.3.6.2 Diverse CDR3β region sequences of Tp1.1 and Tp2.2-specific CTL clones

The sequences of the TCRβ chains expressed by the Tp2.2- and Tp1.1-specific CTL clonotypes are presented in Tables 6.3 and 6.4. There is generally little evidence of conservation of the CDR3β either in terms of length or amino acid sequence within the epitope-specific CTL populations. For example, the two Tp2.2-specific Vβ8s2$^+$ clonotypes show CDR3β with different lengths and no conserved sequence motifs (Table 6.3) and similarly the two large Tp1.1-specific Vβ3s1$^+$ clonotypes (Table 6.4) show no conserved CDR3β features.

However, there are a limited number of examples of CDR3β conservation. Three Tp2.2-specific Vβ28s1$^+$ clonotypes express identical CDR3β regions and Jβ genes (Table 6.3). The two clonotypes in 592 were distinguished on the basis that (i) there was a single nucleotide polymorphism in the sequence of the Vβ28s1 gene, suggesting the expression of different alleles and more significantly (ii) one of the clonotypes co-expressed a non-functional Vβ13s6$^+$ TCRβ chain, which was not expressed by the other clonotype (Appendix L). The expression of identical TCRβ chains by Tp2.2- specific 592 and 1011 clonotypes suggests that this particular Vβ28s1$^+$ chain represents a 'public' TCRβ. Similarly, an identical 'public' Tp1.1-specific Vβ1s9$^+$ TCRβ chain was expressed by a single clonotype in both 468 and a 641 (Table 6.4).

Additionally, amongst the Tp1.1-specific CTL clonotypes there were examples of pairs of Vβ1s7$^+$ and Vβ14s1$^+$clonotypes that exhibited conservation of CDR3β length and partial conservation of the sequence (Table 6.4). In some cases the non-conserved amino acid residues in corresponding positions of these CDR3β regions involved conservative substitutions. The aspartate (D) and glutamate (E) in the 3rd positions of the Vβ1s7$^+$ chains CDR3βs are both polar, hydrophilic and negative residues; and the tryptophan (W) and phenylalanine (F) in the 5th positions of the Vβ1s7$^+$ chains CDR3β chains are both aromatic, hydrophobic and neutral residues.

Sequence					Number of clones in clonotype	Animal
Vβ gene	FR	CDR3	FR	Jβ gene		
1s1	C A S	S Q G R G S G F E Q	Y F G	3s7	2	592
1s6	C A S	S P N S Y E Q	Y F G	3s7	1	592
1s7	C A S	S Q V G G I Y G E L	H F G	3s1	3	1011
2s4	C S A	Q W G G S Y E E Q	H F G	2s1	6	592
2s4	C S A	R F G P G G L S Y E Q	Y F G	3s7	1	1011
4s6	C S A	G V G T R L S Y E Q	Y F G	3s7	4	1011
6s2	C A S	S P I L A W E T L	Y F G	2s4	2	592
8s2	C A S	S K S L F E I Q	Y F G	3s5	2	592
8s2	C A S	S Y E T S G Y G E L	H F G	3s1	1	592
12s1	C A S	H I R G G L D T Q P L	Y F G	3s2	4	592
12s1	C A S	S L V G R E T L	Y F G	2s4	1	1011
12s2	C A S	S Y S P G G G S P L	Y F G	3s3	2	1011
13s1	C A S	S N S R A T A Y D Y	H F G	1s2	1	1011
13s2	C A S	S H A G Y E Q	Y F G	J3	1	592
13s4	C A S	S S S L L D T Q	Y F G	3s4	2	592
13s7	C A S	G H L V G G T R E L	H F G	3s1	5	592
13s8	C A S	S Q G Q Q G Y V N N P L	Y F G	3s3	1	592
13s9	C A S	G E Q	Y F G	3s7	2	592
14s1	C A S	S E S Q G A N Y D Y	H F G	1s2	4	592
14s1	C A S	S V G N S N Y E Q	Y F G	3s7	28	1011
17s1	C A A	S S V F G G D L K S Y E Q	Y F G	3s7	6	592
28s1	C A S	G G R D S I Y D Y	H F G	1s2	6	592
28s1	C A S	A E Y G G E N T Q P L	Y F G	3s2	14	592
28s1	C A S	A E Y G G E N T Q P L	Y F G	3s2	2	592
28s1	C A S	A E Y G G E N T Q P L	Y F G	3s2	1	1011
Xs1	C T C	S K A A A E D G Y E Q	Y F G	3s7	18	1011
Xs1	C T C	S R G D N S Y E Q	Y F G	3s7	1	1011

Table 6.3 TCRβ chain sequence of the Tp2.2-specific CTL clonotypes identified from 592 and 1011. The TCRβ chain sequence is presented in the format of Chothia *et al* (1988). The number of clones represented within each clonotype and the animal from which each clonotype was derived are detailed. The TCRβ chain sequence of the three clonotypes expressing the 'public' Vβ28s1$^+$ TCRβ chain are shaded.

Sequence					Number of clones in clonotype	Animal
Vβ gene	FR	CDR3	FR	Jβ gene		
1s7	C A S	S D D F Y S T D T Q	Y F G	3s4	2	468
1s7	***C A S***	***S H E W Y S T D T Q***	***Y F G***	***3s4***	18	641
1s9	C A S	S Q D Y P T N D P L	Y F G	3s3	8	468
1s9	C A S	S Q D Y P T N D P L	Y F G	3s3	1	641
2s7	C G V	P S G T A S D Y	H F G	1s2	2	468
2s9	C G A	H G P D Y S T D T Q	Y F G	3s4	1	468
3s1	C A S	S Y F G V T Q P L	Y F G	3s2	19	468
3s1	***C A S***	***R E K G L R G S S N D P L***	***Y F G***	***3s3***	34	641
3s1	C A S	N K W G G G N N P L	Y F G	3s3	1	641
4s5	C S A	R P D F G G F N D P L	Y F G	3s3	1	641
4s5	C S A	R S G A W D Y	H F G	1s2	1	641
6s2	C A S	S R S S G H V S N D P L	Y F G	3s3	1	468
10s4	C A S	R R A V M N D P L	Y F G	3s3	17	468
10s7	C A S	S Q R Y D E Q	Y F G	3s7	1	641
13s1	C A S	R G G A G R N G E L	H F G	3s1	1	468
13s1	C A S	R I T A G E Q	H F G	2s1	1	468
14s1	C A S	S A T T A G L A G A A L	T F G	3s6	1	468
14s1	C A S	S S S W E Q	Y F G	3s7	3	641
14s1	C A S	S E S W E L	H F G	2s1	1	641
14s1	C A S	S G G Q L S E R	Y F G	2s5	1	641
14s1	C A S	S D H R A G P W E Q	Y F G	3s7	1	641
16s1	C A S	S L G A L E V	F F G	1s1	1	641
20s1	C A W	S R G N E Q	Y F G	3s7	1	468
24s1	C A S	R A R T P S N D P L	Y F G	3s3	1	468

Table 6.4 TCRβ chain sequence of the Tp1.1-specific CTL clonotypes identified from 468 and 641. The TCRβ chain sequences are presented in the format of Chothia *et al* (1988). The number of clones represented within each clonotype and the animal from which each clonotype was derived are detailed. The TCRβ chain sequence of the clonotypes expressing the 'public' Vβ1s9$^+$ TCRβ chain, or the partially conserved Vβ1s7$^+$ TCRβ chain/Vβ14s1$^+$ TCRβ chain are shaded. The two 641 clonotypes sharing identical TCRβ chain sequence with non-Tp1 specific clonotypes are shown in bold italicised script (compare with Table 6.7)

6.3.7 TCRβ chain sequence of Tp2.1-specific CTL clonotypes

Only three Tp2.1-specific CTL clonotypes were identified from the 592 and 1011 clones. As can be seen from Table 6.5 there was no conservation of either the Vβ genes or CDR3β regions expressed by these clonotypes. However, insufficient numbers of Tp2.2-specific clonotypes have been described to make any definitive observations.

Sequence					Number of clones in clonotype	Animal
Vβ gene	FR	CDR3	FR	Jβ gene		
4s3	CSA	PGQQGYEQ	YFG	3s7	1	1011
13s5	CTS	SRGGRIDGEL	HFG	3s1	1	1011
14s1	CAS	SVSFGGAPYGEL	HFG	3s1	8	592

Table 6.5 TCRβ chain sequence of the Tp2.1-specific CTL clonotypes identified from 592 and 1011. The TCRβ chain sequence is presented in the format of Chothia *et al* (1988). The number of clones represented within each clonotype and the animal from which each clonotype was derived are detailed.

6.3.8 Clonal composition of T-cell clones not cytotoxic for defined epitopes.

The sequences of the non-Tp2.1/Tp2.2-specific CTL clonotypes identified from the 592 and 1011 clones are presented in Table 6.6. This population appears to be extremely polyclonal, with only two clonotypes (a Vβ1s3$^+$ and a Vβ15s2$^+$ clonotype) being represented by more than 1 clone. There is generally little conservation of either Vβ gene or CDR3β expression although there is a pair of Vβ2s3$^+$ clonotypes and a set of four Vβ15s2$^+$ clonotypes that show conserved CDR3β length and partial conservation of CDR3β sequence (Table 6.6 - shaded sequences).

The non-Tp1.1-specific CTL clonotypes present amongst the 468 and 641 clones display a more oligo-clonal nature, with over half (6/11) of the clonotypes being composed of 2 or more clones (Table 6.7). There is no identifiable conservation of Vβ gene or CDR3β region expression by the different non-Tp1.1-specific CTL

clonotypes (Table 6.6). Notably, two non-Tp1.1 cytotoxic 641 clonotypes (the Vβ1s7$^+$ and a Vβ3s1$^+$ clonotype) expressed TCRβ chain sequences identical to those expressed by Tp1.1-cytotoxic 641 clonotypes (compare the bold italicised sequence in Table 6.7 and 6.4).

Sequence					Number of clones in clonotype	Animal
Vβ gene	FR	CDR3	FR	Jβ gene		
1s3	CAS	SQDRRDSYVQL	YFG	2s2	2	592
1s3	CAS	SQGRYNNPL	YFG	3s3	1	592
1s3	CAS	SPAGRDTQ	YFG	3s4	1	1011
1s7	CAS	SQDIGGSTQIQ	YFG	3s5	1	592
1s11	CAS	SSGAANSNNPL	YFG	3s3	1	1011
1s13	CAS	SESRGGGYEQ	YFG	3s7	1	592
2s3	CSA	SMGFRGQSTQ	YFG	2s3	1	1011
2s3	CSA	SWGVGGQSTQ	YFG	2s3	1	1011
5s1	CAS	SPTSGPSNNPL	YFG	3s3	1	1011
6s2	CAS	LASRNTEV	FFG	1s1	1	592
6s2	CAS	SLHGNTEV	FFG	1s1	1	1011
6s3	CAS	SPTGGGSAHNNPL	YFG	3s3	1	1011
8s1	CAS	SKSTDTEV	FFG	1s1	1	592
10s3	CAS	SHRTASYDY	HFG	1s2	1	1011
12s1	CAS	SLVGRETL	YFG	2s4	1	1011
13s1	CAS	SYQGPNPL	YFG	3s3	1	1011
13s2	CAS	SYQPINSPL	HFG	1s5	1	1011
13s4	CAR	GLGYRETL	YFG	2s4	1	1011
15s2	CAS	NSGQQGGTQPL	YFG	3s2	1	1011
15s2	CAS	NPGQQGGTQPL	YFG	3s2	1	1011
15s2	CAS	NPGQQGGAVQL	YFG	2s2	1	1011
15s2	CAS	NRGQQGGSYEQ	YFG	3s7	2	1011
16s1	CAS	SRGEANSETL	YFG	2s4	1	1011
17s2	CSG	ASGLDY	HFG	1s2	1	592
17s3	CAA	SRDMTGTEV	FFG	1s1	1	592
20s1	CAL	RAGGSSETL	YFG	2s4	1	1011

Table 6.6 TCRβ chain sequence of the non-Tp2.1/Tp2.2 specific CTL clonotypes identified from 592 and 1011. The TCRβ chain sequences are presented in the format of Chothia *et al* (1988). The number of clones represented within each clonotype and the animal from which each clonotype was derived are detailed. The TCRβ chain sequence of the clonotypes expressing partially conserved Vβ2s3$^+$ TCRβ chain/Vβ15s2$^+$ TCRβ chains are shaded.

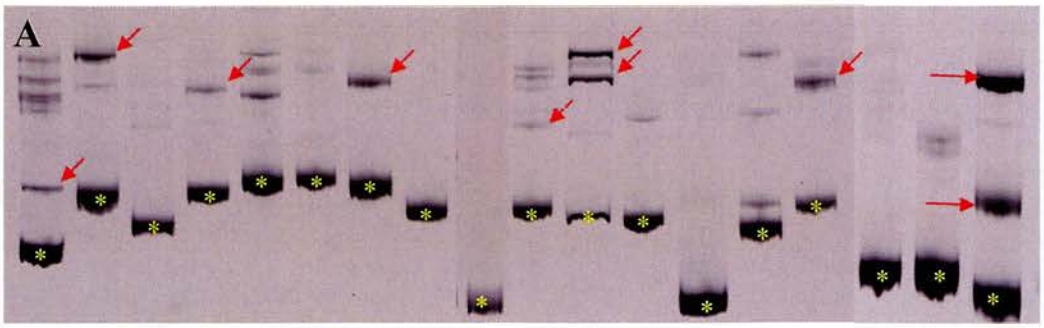
Sequence					Number of clones in clonotype	Animal
Vβ gene	FR	CDR3	FR	Jβ gene		
1s7	***CAS***	***SHEWYSTDTQ***	***YFG***	***3s4***	***4***	***641***
1s9	CAS	SPDIPTIGER	YFG	2s5	2	641
2s7	CGA	RAPRGPSRE	YFG	3s3	2	468
3s1	***CAS***	***REKGLRGSSNDPL***	***YFG***	***3s3***	***9***	***641***
3s1	CAS	SAKAHGRVHAGAAL	TFG	3s6	1	641
7s1	CAS	SREPQDREI	LFG	2s1	1	641
13s2	CAS	SYEEHGEQ	YFG	3s7	7	468
13s10	CAS	RGESHNPL	YFG	3s3	1	641
14s1	CAS	SRGWEQ	YFG	3s7	2	641
14s1	CAS	SKDSVPERL	YFG	3s3	1	468
20s1	CAW	ESEETL	YFG	2s4	1	468
20s1	CAW	THQTESYDY	HFG	1s2	1	468

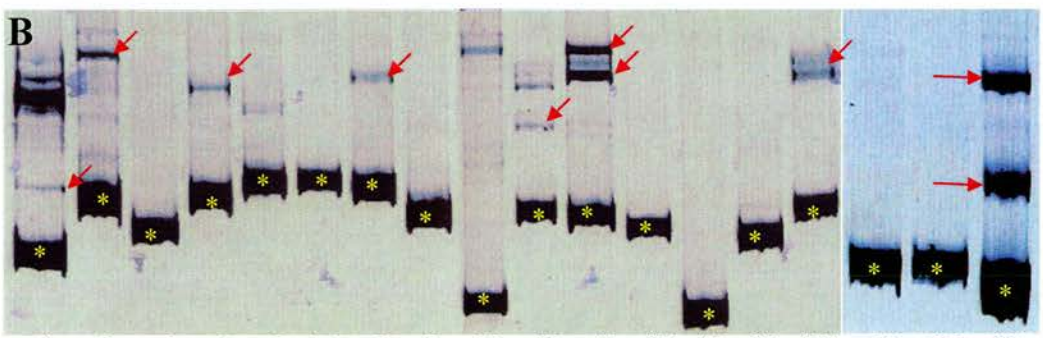
Table 6.7 TCRβ chain sequence of the non-Tp1.1-specific CTL clonotypes identified from 468 and 641 clones. The TCRβ chain sequences are presented in the format of Chothia *et al.* (1988). The number of clones represented within each clonotype and the animal from which each clonotype was derived are detailed. The two 641 clonotypes sharing identical TCRβ chain sequence with Tp1.1-specific CTL clonotypes are shown in bold italicised script (compare with Table 6.4)

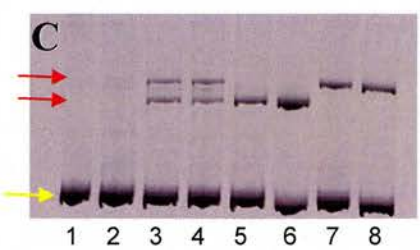
6.3.9 Clonality of the *in vivo* CD8$^+$ T-cell response to *T.parva*

An important question raised by these findings is whether the responses measured *in vitro* are representative of the response of the memory T-cell population following *in vivo* antigenic challenge or reflect biases introduced by *in vitro* manipulation. In Chapter 3, comparative analyses of CD8$^+$ T-cell enriched cell-lines with clones derived from them demonstrated that the CDR3β-HDA technique provides accurate descriptions of the clonal composition of complex T-cell populations (section 3.3.4.2). In order to determine whether the *in vitro* findings reflect the composition of the memory CD8$^+$ T-cell response *in vivo*, one animal (592) was challenged with sporozoites and the clonal composition of CD8$^+$ T-cells harvested at the peak of the response was compared with that measured *in vitro* from PBMC isolated prior to the challenge.

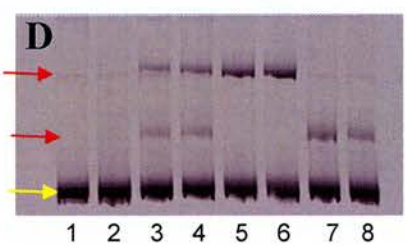
The results of CDR3β-HDA revealed a remarkably similar pattern of clonal composition observed in the *in vivo* response and that measured *in vitro* (Figure 6.9a and 6.9b, red arrows indicate some of the heteroduplex bands evident in both CDR3β-HDA) and co-migration of heteroduplex bands from the *in vivo* response, the response measured *in vitro* and selected representative clones derived from the *in vitro* cell-line could be demonstrated (e.g. Vβ14 in Figure 6.9c and Vβ28 in Figure 6.9d). This result, albeit with one animal, suggests that the responses measured *in vitro* are representative of those seen *in vivo* by *T. parva*-specific memory $CD8^+$ T-cells.







Lane
1- Vβ14 probe + H_2O
2- Vβ14 probe + *ex vivo* pre-challenge
3- Vβ14 probe + *ex vivo* post-challenge
4- Vβ14 probe + *in vitro* response
5- Vβ14 probe + clone 592.3
6- Vβ14 probe + clone 592.4
7- Vβ14 probe + clone 592.26
8- Vβ14 probe + clone 592.66



Lane
1- Vβ28 probe + H_2O
2- Vβ28 probe + *ex vivo* pre-challenge
3- Vβ28 probe + *ex vivo* post-challenge
4- Vβ28 probe + *in vitro* response
5- Vβ28 probe + clone 592.9
6- Vβ28 probe + clone 592.29
7- Vβ28 probe + clone 592.13
8- Vβ28 probe + clone 592.69

Figure 6.8 Comparison of CDR3β-HDA of the *in vitro* and *in vivo* responses of the *T. parva*-specific memory $CD8^+$ T-cell population from animal 592. (A) CDR3β-HDA of the *in vitro* $CD8^+$ T-cell enriched cell-line derived from PBMC of animal 592 taken >2 months after the second *in vivo* challenge. (B) CDR3β-HDA of the *ex vivo* $CD8^+$ T-cell enriched PBMC of animal 592 taken 10d following the third *in vivo* challenge post-immunisation. (C) Comparison of $V\beta14^+$ CDR3β-HDA migration patterns (D) Comparison of $V\beta14^+$ CDR3β-HDA migration patterns. Homoduplices are indicated by yellow asterisk/arrows and red arrows indicated heteroduplex bands. Not all similarities in the CDR3β-HDA results presented in a and b have been indicated.

6.4 Discussion

Results of previous studies have indicated that the protective $CD8^+$ T-cell response against *T. parva* is subject to the phenomenon of immunodominance (Morrison, 1996a; Taracha *et al.*, 1995b), whereby CTL are focused on a limited number of immunogenic epitopes. The results from this chapter, based largely on *in vitro* studies, provide the first direct quantitative evidence of immunodominance in the $CD8^+$ T-cell response to *T. parva* and have yielded detailed information on the clonal composition and TCRβ chain repertoires of CTLs specific for dominant epitopes.

Based on analysis of approximately 90 $CD8^+$ T-cell clones from each of 4 animals immunised against *T. parva* (Muguga), >70% of the clones from each animal were found to be specific for a single dominant epitope, Tp1.1 in the two A18-homozygous animals (468 and 641) and Tp2.2 in the two A10-homozygous animals (592 and 1011). Thus, these epitopes appear to be highly immunodominant in cattle with these particular MHC genotypes immunised against the Muguga strain of *T. parva.* An additional epitope, Tp2.1, recognised by 10% or less of clones from the two A10-homozygous animals appears to be a subdominant epitope.

In contrast to the $CD8^+$ T-cell lines generated from animals 641 and 1011 post-immunisation, those generated pre-immunisation were not cytotoxic for MHC-matched targets presenting the Tp1.1, Tp2.1 or Tp2.2 epitopes(Figure 6.1), indicating that the detectable response to these epitopes was attributable to immunisation. However, both pre- and post- immunisation cell-lines from these 2 animals exhibited cytotoxic activity for the MHC-matched TpM targets. The cytotoxicity displayed by both post-immunisation cell-lines and the 1011 pre-immunisation cell-line appeared to be MHC-restricted whereas the 641 pre-immunisation cell-line also exhibited lysis of the MHC-mismatched TpM target. These findings contradict the reports from previous studies that indicate that *in vitro* stimulated PBMC from naïve animals have cytotoxic activity against allogeneic but not autologous TpM (Eugui and Emery, 1981; Pearson *et al.*, 1979). This discrepancy may be associated with differences between the *in vitro* culture methods used in this study (3 stimulations of PBMC with autologous TpM, the last following enrichment for $CD8^+$ T cells and

supplementation with IL-2) and the studies reported in the literature (single stimulation with autologous TpM, no IL-2 supplementation or enrichment of $CD8^+$ cells).

Phenotypic analysis identified a significant difference in the composition of the cell-lines generated pre- and post- immunisation (Table 6.1). T-cells ($CD3^+$) constituted the majority of the post-immunisation cell-lines populations, whereas in the pre-immunisation cell-lines NK-like cells ($NKp46^+$) were the predominant population. Several $NKp46^+$ clones were established from the pre-immunisation 641 cell-line by limiting dilution and phenotyped as $CD3^-$, $CD4^-$, $CD8^{lo}$ by FACS analysis (data not shown) - consistent with the phenotype of bovine $NKp46^+$ NK-like cells reported in the literature (Storset *et al.*, 2004). Variation in the cytotoxic activity of different bovine $NKp46^+$ populations has been reported (Storset *et al.*, 2004) and may account for the difference between the cytotoxicity profiles of the 1011 and 641 pre-immunisation cell-lines (assuming that the cytotoxicity is due to the $NKp46^+$ populations). NK cells serve as a crucial first line defence against a diverse range of pathogens, including other apicomplexan parasites such as *Plasmodium falciparum* and *Toxoplasma gondii* (Korbel *et al.*, 2004; Lodoen and Lanier, 2006). Although NK-like cells have been identified from the PBMC of a *T. parva*-immunised animal (Goddeeris *et al.*, 1991), the role of these cells is not yet defined and may make an interesting area of investigation in the future.

Sequences for a single predicted functional TCRβ chain were obtained for 310 out of the 355 clones analysed (87%), no TCRβ chains were obtained for 25 of the clones examined (7%) and two predicted functional TCRβ chains were sequenced in the remaining 20 clones (6%). The failure to identify any TCRβ sequence was mainly attributable to loss of mRNA during processing (lack of cDNA demonstrated by failure to amplify the control gene glyceraldehyde-3-phosphate dehydrogenase by PCR - data not shown). For most of the clones with dual functional TCRβ chains, at least one of the expressed chains was identical to a sequence from another clone expressing a single TCRβ, indicating that these clones were actually bi-clonal. As clones selected from plates showing growth in less than 30% of wells seeded by

limiting dilution have only 83% probability of being truly clonal (Henry *et al.*, 1980), the detection of more than one clone in ~6% of the cultures is not unexpected. Alternative explanations for the expression of dual predicted functional TCRβ chains, such as loss of allelic exclusion at the TCRB locus, which has been reported to occur in up to 1% of human peripheral αβ T-cells (Davodeau *et al.*, 1995; Padovan *et al.*, 1995), and the expression of TCRβ chains that although predicted from nucleotide sequence to be functional are in fact non-functional (Rowen *et al.*, 1996) may be applicable but are less likely to be numerically significant.

The response to the immunodominant-epitope in the $CD8^+$ T-cell enriched cell-lines from all 4 animals was found to be polyclonal but dominated by a limited number of large clonal expansions (Figure 6.6). This pattern of clonality has been described for antigen-specific $CD8^+$ T-cell responses to dominant epitopes in numerous immunodominant systems (Annels *et al.*, 2000; Chen *et al.*, 2001c; Maryanski *et al.*, 1999; Naumov *et al.*, 1998). Both random, non-synchronous activation of different antigen-specific clonotypes (Bousso *et al.*, 1999) and the early entry into activation of clonotypes bearing high affinity TCR for the cognate pMHC ligand (Maryanski *et al.*, 1999) have been proposed in other models for the presence of large oligo-clonal expansions amongst the polyclonal response.

Of the 4 animals in this study, only 592 had undergone *in vivo* challenges following immunisation before the response of PBMC was analysed *in vitro*. Interestingly, the Tp2.2-specific CTL clones from 592 were more evenly distributed amongst a larger number of clonotypes than in 468, 641 and 1011, causing a notable 'right-shift' of the 592 curve comparative to the curves of the other 3 animals in Figure 6.6. This may indicate that the number of *in vivo* exposures to *T. parva* has influenced the clonality of the subsequent CTL responses measured *in vitro*, but in the absence of comparative information on the response following the initial immunisation and corroborative data from more animals, this remains speculative. Intriguingly, the immunodominant epitope-specific CTL population of clones was composed of more clonotypes for 592 than for any of the other 3 animals, in contrast to the immuno-focusing of responses (reduction in the number of antigen-specific clonotypes) seen

on repeated antigen exposure in other $CD8^+$ T-cell responses (Busch and Pamer, 1999; Kedzierska *et al.*, 2004; Maryanski *et al.*, 1999; Trautmann *et al.*, 2005; Zhong and Reinherz, 2004). However, this observation may be an artefact based on the number of clones examined from each animal in this study - examination of larger numbers of clones and statistical analysis will be required to validate this.

The TCRβ repertoires expressed by the 121 Tp2.2-specific CTL clones (27 clonotypes) and 119 Tp1.1-specific CTL clones (24 clonotypes) were highly diverse. Within both epitope-specific populations Vβ genes from 11 different subfamilies were used (Figure 6.7) - with 21 and 14 Vβ genes expressed by the Tp2.2- and Tp1.1-specific CTL clonotypes respectively. The CDR3β regions were also characterised by a high degree of diversity (Tables 6.3 and 6.4). Expression of highly diverse TCR repertoires by $CD8^+$ T-cells specific for other immunodominant epitopes has been reported, for example in the responses to the HLA-A11 restricted AVF epitope from EBV (Campos-Lima *et al.*, 1997) and the H2-K^d restricted epitopes LLO 91-99 and p60 217-225 from *Listeria monocytogenes* (Busch and Pamer, 1998; Busch and Pamer, 1999). In both of these models, it has been suggested that the immunodominance of the epitopes correlates with their capacity to elicit a response from a diverse TCR repertoire and the results indicate this may relate to an ability of these epitopes to form highly stable complexes with the presenting MHC molecules (Busch and Pamer, 1998), thereby enabling them to activate T-cells bearing high and low affinity TCR (Busch and Pamer, 1999; Campos-Lima *et al.*, 1997). Further work to analyse the stability of complexes formed by Tp2.2 and Tp1.1 with the respective presenting MHC elements and the affinity of the different antigen-specific TCRs could be informative in this regard.

Another possible contributing factor to the ability of Tp1.1 and Tp2.2 to select diverse TCR repertoires is the conformation of the pMHC ligand they form when bound to the presenting MHCI molecule. Both epitopes are 11mers, - and therefore likely to have a bulged conformation within the MHCI groove. Recent crystallographic studies have shown that immunodominant epitopes from influenza presented by H2-D^b ($PA_{224-233}$) and H2-K^b ($PB1_{703-711}$) have a bulged conformation

(Meijers *et al.*, 2005) and have presented evidence that this feature of the $PA_{224-233}$-D^b complex is fundamental to the selection of a diverse epitope-specific TCR repertoire (Turner *et al.*, 2005).

Diverse TCR repertoire expression by epitope-specific $CD8^+$ T-cells has been correlated with the ability to recognise epitope variants (Campos-Lima *et al.*, 1997; Meyer-Olson *et al.*, 2004) and has been cited as an important factor in the control of 'escape mutants' in viral infections such as HIV (Douek *et al.*, 2002; Lopes *et al.*, 2003) where in many individuals immunodominance focuses the response on a limited number of epitopes at any given time during the infection. Studies using peptides with single amino acid substitutions have demonstrated that T-cell clonotypes expressing different TCRβ chains specific for the same epitope show different patterns of cross-reactivity with substituted variant peptides (Ishikawa *et al.*, 1998; Kim and Kim, 2002; Valmori *et al.*, 1998). Thus, an antigen specific T cell population with a diverse TCR repertoire has an inherently greater potential to recognise variant epitopes than a response with a narrow TCR repertoire. The diverse TCR repertoires expressed by Tp1.1 and Tp2.2-specific CTL would therefore be anticipated to accommodate some degree of cross-recognition of the polymorphic forms expressed by other strains of *T. parva* and so mediate cross-protection. However, as can be seen in Table 6.8, polymorphic versions of Tp1.1 and Tp2.2 show considerable variation, whilst in the studies by Campos-Lima *et al.* (1997), Douek *et al.* (Douek *et al.*, 2002), Lopes *et al.* (2003) and others, that demonstrated cross-recognition of epitope variants, the variant epitopes differed by only one residue. This quantative difference may prove crucial in the ability of Tp1.1- and Tp2.2-specific CTL to recognise epitope variants and offers an interesting area for further exploration.

Epitope	Strain	Amino acid sequence
Tp1	Muguga	VGYPKVKEEML
	Marikebuni	---------II
Tp2.1	Muguga	QSLVCVLMK
	Other	A-IK--AQY A-IK--SHH --IK--SQH
Tp2.2	Muguga	KSSHGMGKVGK
	Other	LT-----RI-R LT------I-R LT-KA-TT---

Table 6.8 Amino acid sequences of epitopes Tp1.1, Tp2.1 and Tp2.2. The amino acid sequences of the epitopes expressed by *T. parva* (Muguga) and other strains are presented in alignment. Dashes represent conserved amino acid residues. Epitope sequence data provided by Dr. E. Taracha and Dr. S. Graham (ILRI, Nairobi, Kenya).

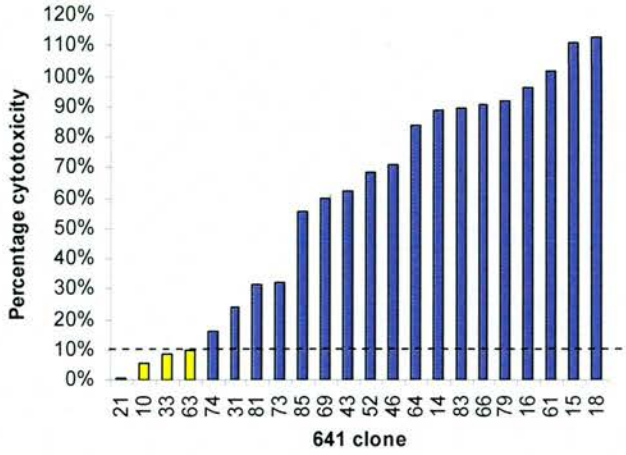
Despite the diversity of the TCRβ repertoires expressed amongst the Tp2.2-specific and Tp1.1-specific CTL populations there were some instances of conservation, including the expression of 'public' TCRβ chains (section 6.3.6.2). It has been suggested that 'public' TCRs represent a 'best fit' for their cognate pMHC ligand and due to their high affinity are preferentially selected from the naïve repertoire and assume progressively greater prominence in the antigen-specific TCR repertoire during chronic antigen exposure or subsequent challenge (Kedzierska *et al.*, 2004; Trautmann *et al.*, 2005). However, both large and small clonotypes expressed the 'public' Vβ28s1^{+} Tp2.2-specific TCRβ chain and Vβ1s9^{+} Tp1.1-specific TCRβ chain (Tables 6.3 and 6.4), indicating that expression of 'public' TCRβ chains was not strictly correlated to large clonal expansion and clonal prominence. Other parameters such as precursor frequencies (Attuil *et al.*, 2000), stochastic selection (Bousso *et al.*, 1999; Kedzierska *et al.*, 2004) and expression of different TCRα chains (Hamrouni *et al.*, 2003) are all likely to be factors in determining clonal abundance. It should be noted that the designation of 'public' TCRβ chain has been applied somewhat loosely in this study - evidence of identical TCRβ chain expression by clonotypes of

the same specificities from more individuals would be required to verify this designation.

A notable feature of the cytotoxicity results obtained in this study was that, with the exception of animal 1011, levels of cytotoxic activity displayed by the clones against the TpM targets were low (Figures 6.2a-6.5a). By contrast, the epitope-loaded MHC-matched Ta targets appeared more sensitive to lysis by CTL, and many clones that demonstrated CTL activity against these targets appeared non-cytotoxic against the TpM target. Thus, the absence of detectable cytotoxicity against the TpM targets in this study is not indicative of a lack cytotoxic capacity of the cells. Since the epitope-loaded Ta targets were pre-incubated with 1μg/ml of peptide, it is likely that the epitope density on the surface of these cells was considerably higher than on the TpM cells. This would imply that the expression of cytotoxic activity is in part dependent on the extent of ligation of TCRs on the T-cells. The generally lower levels of cytotoxicity displayed on TpM targets and the variability between lines may also reflect the inherent susceptibility of the cells to lysis. Different *T. parva*-transformed cells are known to vary in the profiles of cytokines they express and this may be associated with the variation in their susceptibility to CTL lysis (McKeever *et al.*, 1997).

Analyses of TCRβ expression identified a number of Vβ1s7$^+$ (4) and Vβ3s1$^+$ (9) clones that did not give detectable cytotoxicity on either TpM or Tp1.1-loaded A18$^+$ Ta targets but had TCRβ sequences identical to those identified in Tp1.1-specific CTL clonotypes (compare Tables 6.4 and 6.8). The level of cytotoxicity against the Tp1.1-loaded A18$^+$ target exhibited by individual clones sharing either the Vβ1s7$^+$ chain (Figure 6.9a) or the Vβ3s1$^+$ chain (Figure 6.9b) show a gradation ranging from 0.4% to 112.4% for the Vβ1s7$^+$clones and from 15.6% to 96.9% for the Vβ3s1$^+$ clonotypes. This suggests that all of the Vβ1s7$^+$ clones are members of a single clonotype as are all of the Vβ3s1$^+$ clones and that despite the cytotoxicity results, all of these clones are Tp1.1-specific. This finding implies that other clones that were defined as non-Tp2.1/2.2/1.1-specific CTL on the basis of the cytotoxicity results may be also specific for these epitopes. Furthermore, functional heterogeneity of

A



B

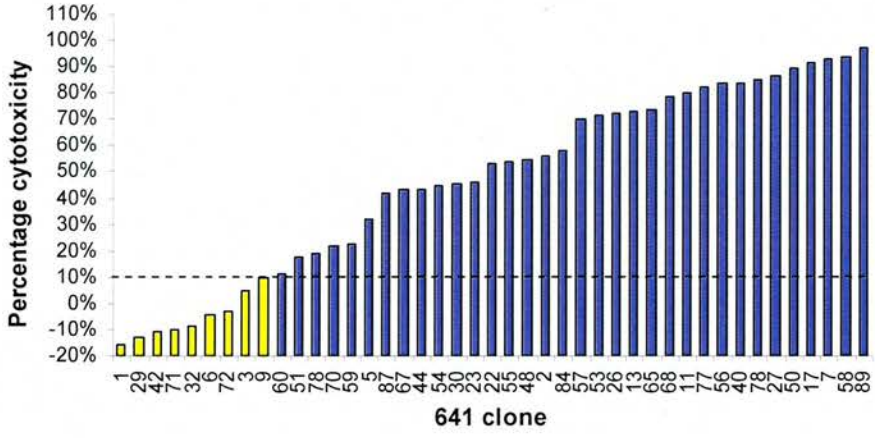


Figure 6.9 Cytotoxic activity against the Tp1-loaded A18^{+}Ta target by 641 clones expressing the (A) Vβ1s7^{+} TCRβ chain (Vβ1s7-SHEWYSTDTQ-Jβ3s4) and (B) Vβ1s7^{+} TCRβ chain (Vβ1s7-REKGLRGSSNDPL-Jβ3s3).The level of activity considered to reflect significant cytotoxicity (≥10%) is represented by the dashed line - clones considered to be non-cytotoxic are presented by yellow bars and clones considered to be cytotoxic by blue bars.

epitope-specific $CD8^+$ T-cells, including cytotoxic and non-cytotoxic subsets, has been reported in responses to viral epitopes (Aandahl *et al.*, 2003; Gillespie *et al.*, 2000; Sandberg *et al.*, 2001). Thus, the percentage figures obtained for the proportions of $CD8^+$ T-cells that are specific for the defined *T. parva* epitopes examined are likely to be underestimated and should be considered as minimum estimates. In future experiments concurrent use of additional functional assays that assess other parameters of T-cell activation such as IFN-γ release, for example using the technique developed by Ballingal *et al.* (2000) or immuno-screening antigen presenting cells (APCs) transiently transfected with *T. parva* (Muguga) schizont cDNA clones (Graham *et al.*, 2006) using ELISpot assays, would provide a more comprehensive analysis of the proportion of clones specific for the defined epitopes.

Amongst the population of T-cell clones that displayed no defined epitope-specific cytotoxic activity were identified a substantial $V\beta13s2^+$ clonotype (Table 6.7) in one of the $A18^+$ animals (468) and sets of $V\beta2s3^+$ and $V\beta15s2^+$ clonotypes with semi-conserved CDR3β regions (Table 6.6) in one of the $A10^+$ animals (1011). This suggests the presence of as yet undefined subdominant epitopes in both the A18- and A10-restricted responses to *T. parva* (Muguga) that are capable of eliciting numerically significant $CD8^+$ T-cell responses. However, it must be noted that as none of these clonotypes displayed cytotoxicity against any target they may be non-cytotoxic cells specific for the already defined epitopes.

As *in vitro* culture of lymphocytes can introduce unknown biases into T-cell responses with regard to the representation of both antigenic specificities (Yewdell and Bennink, 1999) and of TCR composition (Dietrich *et al.*, 1997), results from *in vitro* studies of the $CD8^+$ T-cell response against *T. parva* may not be representative of the *in vivo* response. In the absence of peptide-MHCI tetramers, it is difficult to quantitatively analyse epitope immunodominance directly *ex vivo*. However, comparison of the *in vitro* and *in vivo* responses of the memory $CD8^+$ T-cells in one animal using CDR3β-HDA, demonstrated that the clonal composition of these populations was remarkably similar (Figure 6.9). It could therefore be inferred that the *in vitro* cultures provide a reasonable representation of the *in vivo* memory $CD8^+$

T-cell response and consequently of any immunodominance exhibited. Unfortunately, due to time constraints it was not possible to compare the *in vivo* memory $CD8^+$ T-cell responses and the responses measure *in vitro* for the other animals to corroborate this finding, but hopefully this work will be completed in the future.

In summary the work presented in this study provides the first quantative evidence for immunodominance in the $CD8^+$ T-cell response to *T. parva*. The responses to both the immunodominant A10- and A18-restricted epitopes measured *in vitro* are polyclonal but dominated by a small number of large oligo-clonal expansions and display diverse TCRβ chain repertoires. Furthermore, there is evidence from one animal that there is a high degree of correlation of the composition of the *in vitro* and *in vivo* memory $CD8^+$ T-cell responses, suggesting the immunodominance defined in the *in vitro* studies can be extrapolated to the *in vivo* response. Demonstration of marked immunodominance by polymorphic epitopes in the $CD8^+$ T-cell response to *T. parva* corroborates observations from previous studies and has important implications for the future design of subunit vaccines.

7 Chapter 7 - General discussion

Due to difficulties associated with the application and sustainability of current control measures for *Theileria parva,* current research efforts are directed towards the development of an effective subunit vaccine. As there is evidence that cytotoxic $CD8^+$ T-cell responses against parasitized lymphoblasts have an important role in mediating immunological protection against *T. parva*, CTL target antigens are obvious candidates for inclusion in such a vaccine. Previously described characteristics of the *T. parva*-specific CTL response suggest it is focused on a limited number of immunodominant epitopes that exhibit polymorphism between parasite strains. This immunodominance potentially poses a major obstacle to the development of a *T. parva* subunit vaccine that will effectively elicit CTL with the ability to offer broad protection against the multiplicity of different *T. parva* strains present in the 'field'. In viral systems in which immunodominance has been documented, diversity in the clonality and TCR repertoire of the CTL response has been associated with the ability to recognise epitope variants and thereby constrain 'escape mutants'. Consequently analysis of these two features is of interest when considering potential candidate epitopes for inclusion in subunit vaccines.

The recent identification of CTL target antigens offered the opportunity to study and gain a more detailed understanding of immunodominance in responses to *T. parva*, which will aid the future rational design of a subunit vaccine. The aim of this study was to quantify immunodominance in the *T. parva* $CD8^+$ T-cell response and characterise the clonal composition and TCR repertoire of the responding epitope-specific populations. The work completed in this study had three significant outcomes:

1) Development of molecular tools for analysis of T-cell clonality and the TCRβ chain repertoire of bovine T-cell responses. To achieve the aim of this study it was first necessary to develop the molecular tools to analyse bovine TCRβ chain expression. A panel of Vβ subfamily-specific primers for use in snPCR reactions and a 'Pan-Vβ' snPCR that allowed the amplification of all known bovine TCRβ chains

were developed and validated. Also designed was a CDR3β-HDA that could be used to rapidly and accurately determine the clonal composition of complex bovine αβT-cell populations.

These tools were subsequently used in this study to successfully analyse bovine T-cell responses both *in vitro* and *in vivo*. However, they were designed on the basis of the bovine TCRβ gene sequence data available at the commencement of the study, which has subsequently been expanded by the work within this project. In the light of the expanded gene repertoire identified, some further studies are required to refine and expand this set of tools; for example to accommodate the novel Vβ subfamilies Vβ5, 11 and X and to confirm the specificity of the PCR reagents on the expanded set of Vβ genes. Due to the generic nature of these tools they could be used in future studies to analyze the bovine T-cell response to other pathogens.

2) Extension of the known bovine TCRβ gene repertoire. The large volume of TCRβ chain sequencing completed in the course of this study has identified 38 novel Vβ genes, including members of 4 novel bovine subfamilies (Vβ 5, 12, 25 and X), and substantial numbers of new Vβ10 and Vβ13 subfamily members. Also identified were 4 novel Jβ genes. Examination of the incomplete second assembly of the bovine genome identified a total of 123 Vβ genes, of which 65 were predicted to be functional. Members of five novel bovine subfamilies, Vβ11, 18, 21, 23 and 25, were identified from the genome, although only Vβ11 was predicted to contain a functional member. Although there was substantial similarity in the Vβ repertoires identified from the genome and from cDNA analysis, several genomic genes had no cDNA match and *vice versa*. From the combined cDNA and genomic analysis >90 bovine Vβ which are predicted to be functional have been identified. Partial annotation of the bovine TCRB locus corroborates previous work (Conrad *et al.*, 2002) that suggested the existence of a third DJC cluster and has demonstrated extensive duplication of a cassette containing tandemly located Vβ1 and 13 genes. The bovine Vβ repertoire is now the largest characterised for any species, largely attributable to the prodigious duplication of the Vβ1/13 cassette.

3) Characterisation of immunodominance in the $CD8^+$ T-cell response to *T. parva*. Analysis of the $CD8^+$ T-cell response against *T. parva* (Muguga), utilising *in vitro* derived $CD8^+$ T cell-lines and clones, has shown that the Tp2.2 and Tp1.1 epitopes are highly dominant in animals homozygous for the A10 and A18 class I MHC haplotypes, respectively, with >70% of the T-cell clones evaluated demonstrating specific CTL activity. The responses to the immunodominant epitopes were characterised by a polyclonal response which included a limited number of large clonal expansions and expression of a diverse repertoire of TCRβ chains. For one animal it was possible to demonstrate remarkable similarity between the clonal composition of the responses measured *in vitro* and that of the memory $CD8^+$ T-cell population responding to parasite challenge *in vivo*, inferring that the *in vitro* results give an accurate representation of the *in vivo* response and therefore that the immunodominance noted *in vitro* can be extrapolated to the *in vivo* situation.

A review of the literature indicates that this is the first direct quantative evidence for immunodominance in the $CD8^+$ T-cell response to a parasite and furthermore that it is also the first demonstration of immunodominance in the $CD8^+$ T-cell response to any pathogen in a non-murine/non-primate species. Due to the antigenic complexity of parasites, it might have been predicted that immunodominance would not be an important factor in the $CD8^+$ T-cell responses to this group of pathogens. However, the results described in this study indicate that immunodominance can operate to a similar extent in an anti-parasite $CD8^+$ T-cell response as does in the responses against much less antigenically complex viral pathogens. Additionally, although the TCR repertoires of T-cell response to parasites have been reported previously (e.g. Casanova *et al.*, 1991; Clarencio *et al.*, 2006), the work presented in this thesis represents the most detailed description of such a response so far reported.

However, in many respects the results presented here represent a work in progress. Some of the conclusions reached need to be confirmed and may need to be revised in the light of data from further studies.

The unprecedented expansion of the bovine TCRβ gene repertoire identified from the combined cDNA and genome analysis completed in this study raises many interesting questions regarding the evolutionary history of this locus and the factors that have driven extensive duplication of certain Vβ subfamilies. It is anticipated that the imminent release of the third bovine genome assembly will permit a fuller annotation of the TCRB locus and provide a more exhaustive description of the full bovine TCRβ gene repertoire. Once the sequence of the bovine TCRB locus is complete, comparative analysis of the genomic structure of the bovine, murine and human TCRB loci will offer interesting insights into the evolutionary dynamics of this locus during mammalian speciation. Furthermore, the complete genomic sequence will be useful in examining the enhancer and promoter elements contained within the TCRB locus, that in humans and mice have been shown to influence the naïve TCR repertoire (Chen *et al.*, 2001a; Ryu *et al.*, 2004; Sieh and Chen, 2001) and to be critical in directing the modifications of chromatin structure that control TCRβ rearrangement (Bouvier *et al.*, 1996; Leduc *et al.*, 2000; Oestreich *et al.*, 2006; Sikes *et al.*, 1998). Recent work has made extensive use of cDNA and genomic data to characterise the genomic organisation and repertoire of bovine TCRγ (Herzig *et al.*, 2006a) and TCRδ (Herzig *et al.*, 2006b) genes. However, the bovine TCRα gene repertoire remains poorly characterised, which is unfortunate as knowledge of the TCRα gene repertoire and the subsequent development of tools for analysis of TCRα chain expression would be helpful for unambiguously defining clonotypes within antigen-specific αβT-cell populations.

With evidence for at least 50 serologically defined bovine MHCI specificities (Davies *et al.*, 1994) and a multiplicity of genotypically different *T. parva* strains in the 'field' (Oura *et al.*, 2005), immunodominance of the CTL response by epitopes that exhibit polymorphism between parasite strains may necessitate the incorporation of a vast number of CTL epitopes in the subunit vaccine in order to accommodate all of the pertinent MHC haplotype/parasite strain permutations. In a parallel situation, the combination of immunodominance and diversity of HIV strains and human HLA haplotypes has proved a major obstacle to the successful development of a subunit

vaccine that elicits broadly effective CTL responses against HIV (Altfeld and Allen, 2006; Walker and Korber, 2001).

However, the ability to induce broad protection using relatively few parasite strains in immunisation by infection and treatment (Radley *et al.*, 1975b) suggests that immunodominance may not be an insurmountable obstacle to the design of a broadly effective *T. parva* subunit vaccine. To fully assess the consequences that immunodominance in the CTL response will have on the design on subunit vaccines for *T. parva* it will be necessary to design experiments to address questions such as:

(i) Within defined MHC-restrictions how many epitopes expressed by the global *T. parva* population are capable of dominating the CTL responses? Given the overlapping antigenic profiles of different parasite strains predicted by the mosaic model (section 1.1.3.3), would inclusion of a limited number of immunodominant epitopes sufficiently broaden the specificity of the CTL response on these defined MHC backgrounds to induce protection against the majority of parasite strains?

(ii) Does the diverse clonal composition and TCRβ repertoire of the CTL specific for immunodominant epitopes influence the cross-reactivity of the CTL response and therefore the capacity to offer protection against heterologous parasite strains? In viral responses diverse clonal composition and TCR repertoires are associated with recognition of epitope variants and constraint of the 'escape mutants' that express them (Lopes *et al.*, 2003; Meyer-Olson *et al.*, 2004). There are now established broad panels of Tp2.2 (Muguga)- and Tp1.1 (Muguga)-specific clonotypes whose ability to recognise the epitope variants expressed by other parasite strains could be directly evaluated using *in vitro* assays. Results from such assays could begin to address this question.

(iii) Does *T. parva* express conserved subdominant CTL epitopes which could serve as targets for a subunit vaccine that will induce CTL against a broad spectrum of parasite strains? In experimental murine viral systems subdominant epitope-specific CTL have been shown to afford protection against *in vivo* challenge (Chen *et al.*, 1998; Oukka *et al.*, 1996) and the ability to generate CTL responses to subdominant

epitopes by vaccination has been shown in both primate and murine models (e.g. Chen *et al.*, 1998; Cole *et al.*, 1997; Liu *et al.*, 2006; Santra *et al.*, 2002). Furthermore, effective control of HIV has been correlated to $CD8^+$ T-cell response against subdominant rather than immunodominant epitopes (Frahm *et al.*, 2006) and consequently a current strategy being employed in HIV subunit vaccine design is the identification and targeting of epitopes on the basis of their degree of conservation rather than their position in the immunodominance hierarchy (Altfeld and Allen, 2006).

The results from studies answering these questions will provide significant information that could have profound influence on future strategies employed in the design of a *T. parva*-specific subunit vaccine.

Studies extending the observations made in the A10- and A18- restricted responses to *T. parva* (Muguga) to other combinations of bovine MHCI haplotypes and parasite strains would confirm that immunodominance is a general feature of the $CD8^+$ T-cell response against *T. parva.* Already, studies on immunodominance in the A14-restricted response to *T. parva* (Muguga) are being undertaken by other members of the group. However, it would also be of interest to undertake further studies on the responses described in this study. Application of the molecular tools developed in this study could be used to address issues relating to (i) the stability of the repertoire of epitope-specific CTL clonotypes and the persistence of individual clonal expansions, (ii) the extent of immuno-focusing that occurs during repeated *in vivo* parasite challenge and (iii) the kinetics of development of the $CD8^+$ T-cell response during primary immunisation or infection. The data from such studies would help assess whether these features may impact on the strain-specificity of induced CTL responses. Development and use of additional tools such as real-time CDR3β-specific PCR amplification (Gallard *et al.*, 2002; Hohlfeld, 2003) could add another dimension by enabling examination of the kinetics of individual epitope-specific clonotypes expansion/contraction during the response to *T. parva*, whilst the generation and validation of relevant pMHC-tetramers (currently being completed) will provide a means of directly quantifying immunodominance in the *in vivo*

response to *T. parva* and isolating *ex vivo* epitope-specific T-cells for analysis of their function, clonal composition and expressed TCR repertoire.

The quantitative data on the immunodominance of the Tp2.2 and Tp1.1 epitopes in responses of $A10^+$ and $A18^+$ animals provide a basis for further studies to examine the mechanisms that establish the immunodominance of the Tp2.2 and Tp1.1 epitopes in these responses. Using materials generated in this study (e.g. epitope-specific CTL clonotypes) and other available materials (e.g. the $TAP^{-/-}$ RMA-S cell line transfected with $A10^+$ MHC products) it is now possible to devise experimental assays to study factors known to influence immunodominance in other systems such as, the stability of pMHC complexes (Busch and Pamer, 1998; Chen *et al.*, 1994; Gallimore *et al.*, 1998) and the ability to elicit both high and low avidity $CD8^+$ T-cells (Busch and Pamer, 1998; Campos-Lima *et al.*, 1997).

In summary, the results presented in this study have quantitatively demonstrated immunodominance in the $CD8^+$ T-cell response to *T. parva* and characterised the clonal composition and TCRβ repertoire of the responding T-cells. Evidence of immunodominance against polymorphic epitopes in an outbred population poses potential obstacles for the development of a subunit vaccine against this pathogen. Further investigations will be necessary to evaluate the consequences this has on the selection of candidate CTL epitopes to be included in any future *T. parva* subunit vaccine.

8 Appendices

Appendix A - Correspondence between the IMGT and WHO-IUIS nomenclature systems for human TCRβV

The following table showing the corresponding names of human TCRβV genes according to the IMGT and WHO-IUIS nomenclature systems is derived from the IMGT website (http://imgt.cines.fr). Under the IMGT and WHO-IUIS systems these genes would be prefixed with TRBV and TCRβV respectively. The genes are presented in the 3' to 5' order in which they are arranged in the TCRB locus.

IMGT Nomenclature	WHO-IUIS Nomenclature	IMGT Nomenclature	WHO-IUIS Nomenclature	IMGT Nomenclature	WHO-IUIS Nomenclature
30	20s1	13	23s1	10-1	12s2
29-1	4s1	7-9	6s4	9	1s1
28	3s1	5-8	5a4	5-3	5s5
27	14s1	7-8	6s2	8-2	30s2
26	28s1	6-9	13s4	7-3	6s1
25-1	11s1	5-7	5s7	6-4	13s5
24-1	15s1	7-7	6s6	5-2	5s8
23-1	19s1	6-8	13s7	8-1	30s1
22	29s1	5-6	5s2	7-2	6s5
21-1	10s1	7-6	6s3	6-3	13s2b
20-1	2s1	6-7	13s8	4-3	7s2
19	17s1	5-5	5s3	3-2	9s2
18	18s1	7-5	6s9	6-2	13s2a
17	26s1	6-6	13s6	4-2	7s3
16	25s1	5-4	5s6	7-1	6s7
15	24s1	7-4	6s8	6-1	13s3
14	16s1	6-5	13s1	5-1	5s1
12-5	8s3	12-2	8s5	4-1	7s1
12-4	8s2	11-2	21s3	3-1	9s1
12-3	8s1	10-2	12s3	2	22s1
11-3	21s2	12-1	8s4	1	27s1
10-3	12s1	11-1	21s1		

Appendix B - GenBank accession numbers

All deposited sequences can be obtained from the NCBI nucleotide database (http://www.ncbi.nlm.nih.gov)

(i) Genomic sequence of Human TCR Loci

AE000658-AE000662 *Homo sapiens* TCRA/D locus submitted by Koop *et al.*(1994)

U66059-U66061 *Homo sapiens* TCRB locus submitted by Rowen *et al.* (1996)

AF029308 *Homo sapiens* partial duplication of TCRB locus on chromosome 9 - direct submission by Rowen, L. *et al.*

AF159056 *Homo sapiens* TCRG locus - direct submission by Zhan, M. *et al.*

(ii) Genomic Sequence of Murine TCR Loci

AE007512 *Mus musculus* TCRA/D locus - direct submission by Lee, I.Y. *et al.*

AE000663-AE000665 *Mus musculus* TCRB locus submitted by Chen *et al.* (2001a)

AF037352 *Mus musculus* TCRG locus - gamma 1 and 3 clusters - direct submission by Zhou, Q. *et al.*

AF021335 *Mus musculus* TCRG locus - gamma 2 and 4 clusters - direct submission by Parlee, M.J. *et al.*

(iii) Genomic sequence of the bovine TCR loci

AF453325 Partial sequence of the TCRB locus Conrad *et al.* (2002)

AY644517 Bos taurus T cell receptor gamma cluster 1 (TCRG1) - direct submission by Conrad M.L. *et al* .

AY644518 Bos taurus T cell receptor gamma cluster 2 (TCRG2) - direct submission by Conrad M.L. *et al* .

(iv) Genomic sequence of the porcine TCR loci

AB079894	Partial sequence of the TCRB locus - direct submission by Takagaki, Y. *et al.*

(v) Bovine TCR genes

D90121-D90130	Bovine TCRβV gene sequences submitted by Tanaka *et al.* (1990)
L18951	Bovine Vβ6s2 gene sequence submitted by Buitkamp *et al.* (1993)
AJ006346-AJ006347	Bovine TCRβV gene sequences submitted by Houston *et al.* (1997)
AJ006567-AJ006583	Bovine TCRβV gene sequences submitted by Houston *et al.* (1997)
AJ235264-AJ235268	Bovine TCRβV gene sequences submitted by Houston *et al.* (1997)
D90139-D90140	Bovine TCRβC gene sequences submitted by Tanaka *et al.* (1990)

(vi) Human nesprin-1

AF495910	sequence submitted by Zhang *et al.* (2002)

(vii) Porcine TCR genes

AY690918	Porcine VβX gene sequence submitted by Butler *et al.* (2005)

Appendix C – Solutions and media

All reagents were obtained from Sigma (Poole, Doreset, UK) unless otherwise stated. Double-distilled water (DDW) was sourced from a Millipore-RO 60 Plus unit (Millipore, Billerica, MA, USA).

C.1 Alsever's Solution 10x stock

D-glucose	205.0 g	(113.8mM)
Citric acid	5.5 g	(2.9mM)
Sodium chloride	42.0 g	(71.9mM)
Tris-sodium citrate Di-hydrate	80.0 g	(27.2mM)

Made up to a volume of 1 L with DDW.

Made up to 1x solution by filter sterilizing 0.1 L stock solution using 0.45 μm Minisart single-use filter (Sartorius, Goettingen, Germany) and adding to 0.9 L of sterile DDW.

C.2 RBC Lysis Buffer

Trizma base	10.6 g	(0.175M)

0.175M TRIS pH adjusted to 7.4 with concentrated HCl and made up to a volume of 0.5L with DDW

Ammonium chloride	8.55 g	(0.16M)

Made up to a volume of 1L with DDW.

RBC lysis buffer consists of 1 part 0.175M TRIS pH 7.4 and 9 parts 0.16M ammonium chloride. Filter sterilized using 0.45 μm Minisart single-use filter (Sartorius, Goettingen, Germany)

C.3 Standard Culture Medium (SCM)

RPMI 1640 medium + 25mM HEPES + L-Glutamine (Gibco, Paisley, UK)

10% Foetal bovine serum (Gibco, Paisley, UK)

100 U/ml Penicillin, 100 μg/ml streptomycin, 292 μg/ml L-glutamine

(Penicillin-Streptomycin-Glutamine (100x) solution (Gibco, Paisley, UK)

2-Mercaptoethanol (2-ME) 5 $x10^{-4}$M

C.4 FACS Medium

RPMI 1640 medium + 25mM HEPES + L-Glutamine (Gibco, Paisley, UK)

2% Foetal bovine serum (Gibco, Paisley, UK)

0.2% Sodium Azide

C.5 Cytotoxicity Medium

RPMI 1640 medium + 25mM HEPES + L-Glutamine (Gibco, Paisley, UK)

5% Foetal bovine serum (Gibco, Paisley, UK)

C.6 Phosphate Buffered Saline (PBS)

Sodium chloride	170.0 mM
Potassium chloride	3.4 mM
Sodium phosphate, dibasic	9.2 mM
Potassium phosphate, monobasic	1.8mM

pH6.8

PBS tablets used according to manufacturer's instructions (Unipath LTD, Basingstoke, Hampshire, UK)

C.7 SM-0005 PCR Buffer

45 mM Tris-HCl (pH 8.8 at 25°C)

11 mM Ammonium sulphate

4.5 mM Magnesium chloride

0.113 mg/ml Bovine serum albumin

4.4 µM EDTA

1.0 mM each of dATP, dCTP, dGTP and dTTP

Purchased from ABgene (Epsom, Surrey, UK)

C.8 Loading Buffer for Agarose Gel Electrophoresis

15% w/v Ficoll (Type 400) in DDW with bromophenol blue (0.25%v/v) and xylene cyanol (0.25%) dyes

C.9 Tris-acetate/EDTA electrophoresis buffer (TAE) 50x stock

Trizma base	242 g	(2M)
Glacial acetic acid	57.1 ml	(2M)
EDTA	100 ml of 0.5M solution	(50mM)

pH adjusted to 7.7-8.0 with glacial acetic acid if required, made up to a volume of 1 L with distilled water

C.10 SOC Media

Tryptone	20.00 g
Yeast extract	5.00 g
Sodium chloride	0.50 g
Potassium chloride	0.19 g
Glucose	4.00 g

pH adjusted to 7.0 with 5M sodium hydroxide, made up to a volume of 1 L with DDW and autoclaved

C.11 Luria-Bertani (LB) Agar plates with ampicillin

Sodium chloride	10 g
Yeast extract	5 g
Tryptone	10 g
Agar	15 g

pH adjusted to 7.0 with 5M sodium hydroxide, made up to a volume of 1 L with DDW and autoclaved

Upon use, LB agar was melted and ampicillin added to give a final concentration of 100 µg/ml. The required amount of media was poured into sterile Petri dishes and allowed to set and dry. If LB/ampicillin/IPTG/X-Gal plates were required for blue/white colony selection 100 µl of 24 mg/ml IPTG (Isopropyl β-D-1-thiogalactopyranoside) and 50 µl of 50 mg/ml XGal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) in 50 µl of SOC media (Appendix C.10) was applied to the surface of the media and allowed to dry before application of any culture.

C.12 LB Media

Sodium chloride	10 g
Yeast Extract	5 g
Tryptone	10 g
Agar	15 g

pH adjusted to 7.0 with 5M sodium hydroxide, made up to a volume of 1 L with DDW and autoclaved

C.13a Cell Re-suspension Solution for Plasmid Mini-Preparations (Method A)

Tris-HCl pH7.5	5 ml of 1M solution	(50mM)
EDTA	2 ml of 0.5M solution	(100mM)

Made up to a volume of 0.1 L with DDW and autoclaved. RNAase added at 100 µg/ml prior to use.

C.13b Cell Lysis Solution for Plasmid Mini-Preparations (Method A)

Sodium hydroxide	50 ml of 0.4M solution	(0.2M)
Sodium lauryl sulphate (SDS)	50 ml of 2% solution	(1%w/v)

C.13c Neutralisation Solution for Plasmid Mini-Preparations (Method A)

Potassium acetate	12.55 g	(1.32M)

pH adjusted to 4.8 with glacial acetic acid, made up to 0.1 L with DDW and autoclaved.

C.14 Loading Buffer for Polyacrylamide Gel Electrophoresis

40% w/v sucrose in distilled water with bromophenol blue (0.25%v/v) and xylene cyanol (0.25%) dyes

C.15 9% non-denaturing polyacrylamide gel preparation

Enough for two Biorad Protean ® II xi system gels (80 mls)

30% acryl/bis-arcrylamide 24 ml

10 x TBE	4 ml
DDW	52 ml

96 µl TEMED (N, N,N', N-Tetramethylethylene-diamine), and 184 µl 25% ammonium persulphate (1 g ammonium persulphate in 4 ml of DDW. 200 µl aliquots stored at –20°C.) added for polymerisation.

C.16 Tris-borate/EDTA eletrophoresis buffer (TBE) 10x stock

Trizma base	128 g	(890mM)
Orthoboric acid	55 g	(890mM)
EDTA	40 ml of 0.5M solution	(20mM)

Made up to a volume of 1 L in distilled water.

C.17 20%w/v Polyethylene glycol (PEG)/2.5M NaCl

Sodium chloride	58.44 g	(2.5M)
PEG MW8000	100.00 g	(20%w/v)

Made up to a volume of 0.5 L in DDW.

C.18 5 x CSA Buffer

Trizma base	4.840 g	(400mM)
Magnesium chloride	0.203 g	(10mM)

pH adjusted to 9.0 with concentrated HCl, made up to a volume of 0.1 L with DDW and autoclaved.

C.19 3M Sodium Acetate pH 5.2

Sodium acetate.$3H_20$	40.81 g	(3M)

pH adjusted to 5.2 with glacial acetic acid, made up to a final volume of 0.1 L in DDW and autoclaved.

C.20 Supplemented D-MEM medium

D-MEM + 4500mg/L Glucose + L-glutamine + Pyruvate (Gibco, Paisley, UK)

10% Foetal bovine serum (Gibco, Paisley, UK)

100U/ml Penicillin, 100µg/ml streptomycin, 292µg/ml L-glutamine (Penicillin-Streptomycin-Glutamine (100x) solution (Gibco, Paisley, UK)

C.21 Depurination Solution – 0.125M HCl (Southern Blot)

Hydrochloric acid	11 ml concentrated HCl (11M)	(0.125M)

Made up to a volume of 1L with distilled water.

C.22 Denaturation Solution - 1.5M NaCl, 0.5N NaOH (Southern Blot)

Sodium chloride	87.66g	(1.5M)
Sodium hydroxide	20.00g	(0.5M)

Made up to a volume of 1L with distilled water.

C.23 Neutralization Solution - 1M Tris (pH7.4), 1.5M NaCl (Southern Blot)

Sodium chloride	87.66g	(1.5M)
Trizma base	60.50g	(1M)

pH adjusted to 7.5 with concentrated HCl and made up to a volume of 1L with distilled water.

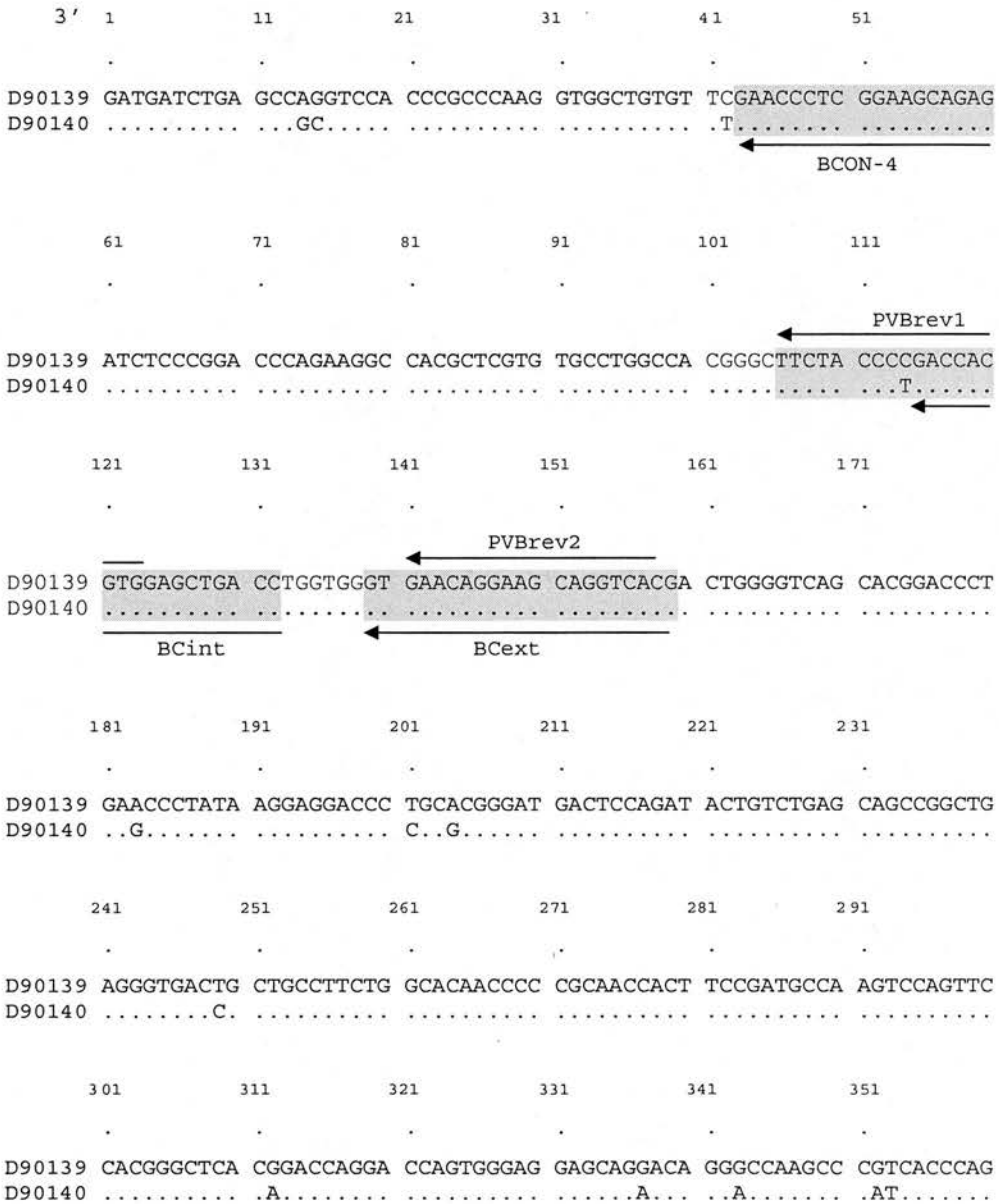
C.24 20 x SSC – 3M NaCl, 0.3M Na_3citrate (Southern Blot)

Tris-sodium citrate	88.23g	(0.3M)
Sodium chloride	175.32g	(3M)

Made up to a volume of IL with distilled water.

Appendix D - Annealing sites of the 3' TCRβC-specific primers

The annealing sites of the 5 different 3' TCRβC-specific primers used in this study are shown on the 2 deposited bovine TCRβC gene sequences (GenBank accession numbers D90139 and D90140 - Appendix B)



```
      361         371         381         391         401         411
       .           .           .           .           .           .
D90139 AACATCAGCG CCGAGGCCTG GGGCAGAGCA GACTGTGGTG TCACCTCCGC ATCTTATCAG
D90140 .......... .......... .......... ........C. .......... G.........

      421         431         441         451         461         471
       .           .           .           .           .           .
D90139 CAAGGCGTCC TGTCTGCCAC CCTCCTCTAT GAGATCCTGC TGGGGAAGGC CACCCTGTAC
D90140 .....G.... .......... ...G...... .......... .......... ..........

      481         491         501         511         521         531        5'
       .           .           .           .           .           .
D90139 GCTGTGCTGG TCAGCGCCCT GGTGCTAATG GCCATGGTCA AGAGAAAAGA GTCCTGAAGC
D90140 .......... .......... ......G... ........T. ...A...G.. T.........
```

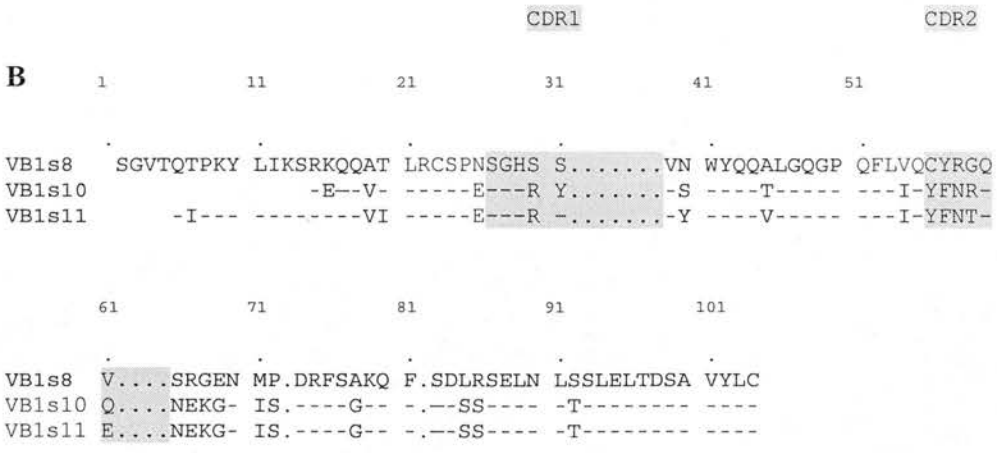
Appendix E - Additional bovine Vβ gene sequences

Sequence data for bovine Vβ genes Vβ1s8, Vβ1s10, Vβ1s11, Vβ2s2, Vβ2s5, Vβ2s7, Vβ4s2, Vβ6s3, Vβ7s2, Vβ9s1 and Vβ17s2 were obtained from Houston (1997) and for Vβ4s3, Vβ4s4, Vβ8s1, Vβ8s2 and Vβ17s2 from data made available by Russell and MacHugh (unpublished). Sequences have been aligned and CDR1 and CDR2 (shaded regions) defined according to the IMGT unique numbering system (Giudicelli *et al.*, 2004). Identity is shown as dashes and gaps by dots.

E.1 Nucleotide (A) and amino acid (B) sequence of Vβ1s8, Vβ1s10 and Vβ1s11

A

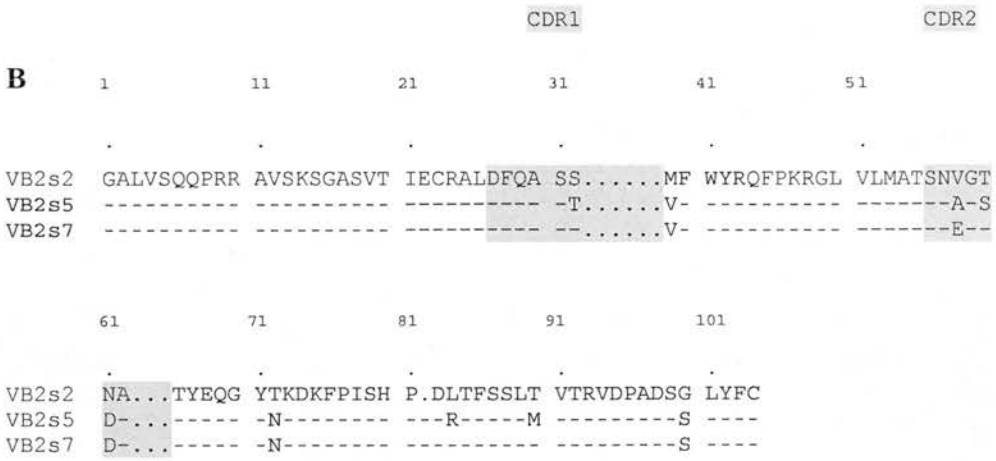
```
          1          11         21         31         41         51
          .          .          .          .          .          .
VB1s8     ..TTCTGGAG TCACCCAGAC CCCCAAATAC CTGATCAAAT CAAGAAAGCA GCAAGCGACT
VB1s10                                                ---G---- -----T----
VB1s11                ----T ---------- ---------- ---------- -----T--TA

                                          CDR1
         61         71         81         91        101        111
          .          .          .          .          .          .
VB1s8     CTGAGATGTT CCCCTAACTC TGGACACAGC TCT....... .......... ....GTGAAC
VB1s10    ---------- ----CG-A-- -------C-- ---....... .......... ....--ATC-
VB1s11    ---------- -----G-A-- -------C-- -A-....... .......... ....---T--

                                                                   CDR2
        121        131        141        151        161        171
          .          .          .          .          .          .
VB1s8     TGGTACCAAC AGGCCCTGGG CCAGGGCCCC CAGTTCCTCG TTCAGTGTTA CCGCGGACAA
VB1s10    ---------- --A------- ------T--- --A------A ------A--T -AA-A-G---
VB1s11    ---------- ---T------ ---------- --A------A ------A--T -AA-ACG---

        181        191        201        211        221        231
          .          .          .          .          .          .
VB1s8     GTT....... .....AGTAG AGGAGAAAAC ATGCCA...G ATCGATTCTC AGCAAAACAA
VB1s10    CAG....... .....-A-GA GAA--G---- --AT--...- ---------- --G------G
VB1s11    -AG....... .....-A-GA GAA--G---- --AT--...- ---------- --G------G


        241        251        261        271        281        291
          .          .          .          .          .          .
VB1s8     TTT...AGTG ACTTGCGCTC TGAGCTGAAT TTGAGCTCCT TGGAGCTGAC AGACTCAGCC
VB1s10    -C...----- ---CTA---- ---------C ----CT---- ---------- G---------
VB1s11    -C...----- ---CTA---- ---------C ----CT---- ---------- G---------


        301        311
          .          .
VB1s8     GTGTATCTCT GT
VB1s10    ---------- --
VB1s11    ---------- --
```


CDR1
CDR2
B
1
11
21
31
41
51
VB1s8 SGVTQTPKY LIKSRKQQAT LRCSPNSGHS S.....VN WYQQALGQGP QFLVQCYRGQ
VB1s10 -E--V- -----E---R Y.....-S ----T----- ---I-YFNR-
VB1s11 -I--- --------VI -----E---R -.....-Y ----V----- ---I-YFNT-
61
71
81
91
101
VB1s8 V....SRGEN MP.DRFSAKQ F.SDLRSELN LSSLELTDSA VYLC
VB1s10 Q....NEKG- IS.----G-- -.--SS---- -T-------- ----
VB1s11 E....NEKG- IS.----G-- -.--SS---- -T-------- ----

E.2 Nucleotide (A) and amino acid (B) sequence of Vβ2s2, Vβ2s5 and Vβ2s7

A

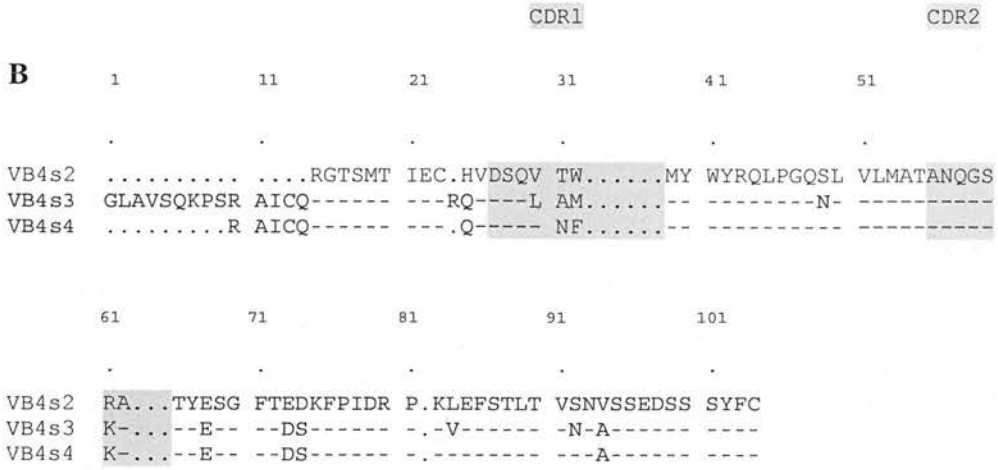
```
          1          11         21         31         41         51
          .          .          .          .          .          .
VB2s2  GGCGCTCTCG TCTCTCAGCA GCCCCGCAGG GCTGTCAGTA AGAGTGGAGC CTCTGTGACC
VB2s5  ---------- ---------- ---------- ---------- ---------- ----------
VB2s7  ---------- ---------- ---------- ---------- ---------- ----------
                                        CDR1

          61         71         81         91        101        111
          .          .          .          .          .          .
VB2s2  ATCGAGTGCC GTGCACTGGA CTTTCAAGCC TCAAGT.... .......... ....ATGTTC
VB2s5  ---------- ---------- ---------- ----C-.... .......... ....G-----
VB2s7  ---------- ---------- ---------- ------.... .......... ....G-----
                                                              CDR2

         121        131        141        151        161        171
          .          .          .          .          .          .
VB2s2  TGGTACCGTC AGTTCCCGAA ACGAGGCCTC GTGCTGATGG CAACTTCTAA TGTGGGCACT
VB2s5  ---------- ---------- ---------- ---------- ---------- --C-----G-
VB2s7  ---------- ---------- ---------- ---------- ---------- --A-------

         181        191        201        211        221        231
          .          .          .          .          .          .
VB2s2  AATGCC.... .....ACATA CGAACAAGGT TATACCAAGG ACAAGTTTCC CATCAGTCAT
VB2s5  G-----.... .....----- ---------- ----A----- ---------- ----------
VB2s7  G-----.... .....----- ---------- ----A----- ---------- ----------

         241        251        261        271        281        291
          .          .          .          .          .          .
VB2s2  CCA...GACC TAACATTTTC ATCTCTGACG GTGACACGTG TGGATCCTGC AGACAGCGGC
VB2s5  ---...---- G--------- --------T- ---------- ---------- -------A--
VB2s7  ---...---- ---------- ---------- ---------- ---------- -------A--

         301        311
          .          .
VB2s2  CTCTACTTCT GC
VB2s5  ---------- --
VB2s7  ---------- --
```


CDR1
CDR2
B
1
11
21
31
41
51
VB2s2 GALVSQQPRR AVSKSGASVT IECRALDFQA SS.....MF WYRQFPKRGL VLMATSNVGT
VB2s5 ---------- ---------- ---------- -T.....V- ---------- -------A-S
VB2s7 ---------- ---------- ---------- --.....V- ---------- -------E--
61
71
81
91
101
VB2s2 NA...TYEQG YTKDKFPISH P.DLTFSSLT VTRVDPADSG LYFC
VB2s5 D-...----- -N-------- ---R-----M ---------S ----
VB2s7 D-...----- -N-------- ---------- ---------S ----

E.3 Nucleotide (A) and amino acid (B) sequence of Vβ4s2, Vβ4s3 and Vβ4s4

A

```
          1           11          21          31          41          51
          .           .           .           .           .           .
VB4s2  .......... .......... .......... .......... ..CGCGGGAC CTCCATGACG
VB4s3  GGTGCTCTCG TCTCTCAAAA GCCAAGCAGG GCCATCTGTC AA--T----- ----G---T-
VB4s4  .......... .......... .....GCAGG GCCATCTGTC AA-------- --------T-
                                          CDR1

          61          71          81          91          101         111
          .           .           .           .           .           .
VB4s2  ATCGAGTGTC ...ATGTCGA TAGCCAAGTC ACCTGG.... .......... ....ATGTAC
VB4s3  ---------- GTC-G--T-- ------GC-- G--A--.... .......... ....------
VB4s4  ---------- ...-G----- ------G--- -AT--C.... .......... ....------
                                                                  CDR2

         121         131         141         151         161         171
          .           .           .           .           .           .
VB4s2  TGGTACCGAC AGCTTCCAGG ACAGAGCTTG GTGTTGATGG CTACTGCCAA TCAGGGCTCC
VB4s3  --------T- ---------- -----A---- ---C------ ---------- ----------
VB4s4  --------T- ---------- ---------- ---C------ ---------- ----------

         181         191         201         211         221         231
          .           .           .           .           .           .
VB4s2  AGGGCT.... .....ACTTA CGAGAGTGGG TTTACCGAGG ACAAGTTTCC CATTGACCGC
VB4s3  -A----.... .....----- ---------- -----T--CA G--------- ----------
VB4s4  -A----.... .....----- ---------- -----T--CA G--------- ----------

         241         251         261         271         281         291
          .           .           .           .           .           .
VB4s2  CCG...AAAC TGGAGTTCTC AACTCTGACT GTGAGCAACG TGAGCTCCGA AGACAGTAGC
VB4s3  ---...---G ---------- ---------- ----A---T- C--------- ---T------
VB4s4  ---...---- ---------- ---------- ---------- C--------- ------C---

         301         311
          .           .
VB4s2  TCTTATTTCT GC
VB4s3  ---------- --
VB4s4  ---------- --
```


CDR1
CDR2
B
1
11
21
31
41
51
VB4s2 ....RGTSMT IEC.HVDSQV TW.....MY WYRQLPGQSL VLMATANQGS
VB4s3 GLAVSQKPSR AICQ------ ---RQ----L AM.....-- --------N- ----------
VB4s4 .....R AICQ------ ---.Q----- NF.....-- ---------- ----------
61
71
81
91
101
VB4s2 RA...TYESG FTEDKFPIDR P.KLEFSTLT VSNVSSEDSS SYFC
VB4s3 K-...--E-- --DS------ -.-V------ -N-A------ ----
VB4s4 K-...--E-- --DS------ -.-------- ---A------ ----

E.4 Nucleotide (A) and amino acid (B) sequence of Vβ8s1 and Vβ8s2

A

```
          1          11          21          31          41          51
          .           .           .           .           .           .
VB8s1  AACCTAGGAG TCACCCAGAC CCCTAGGCAT GAGGTGACAG AAAAGGGACA GGCAGTCACT
VB8s2  ......---- ---------- ---------- ---------- ---------- ---G------
                                          CDR1

          61          71          81          91         101         111
          .           .           .           .           .           .
VB8s1  CTGAGTTGTG AGCTGATTAA AAGCCACACT GCC....... .......... ....CTTTAT
VB8s2  ---------- ---C------ -------G-- ---....... .......... ....----T-
                                                                   CDR2

         121         131         141         151         161         171
          .           .           .           .           .           .
VB8s1  TGGTACAGAC AGACCTCGGT GCGGGGACTG GAGTTTCTGG CTTACTTCAG CAACCAAGCT
VB8s2  ---------- ---------- ---------- ---C-G---- T---T----A ----------

         181         191         201         211         221         231
          .           .           .           .           .           .
VB8s1  CCT....... .....ATTGA TGAGTCGGGG ATGCCCAAGG ACCGATTCTC AGCTAAAATG
VB8s2  ---....... .....----- ---------- ---------- ---------- ------G---

         241         251         261         271         281         291
          .           .           .           .           .           .
VB8s1  CCT...AACT CGTCGTTCTC CACTCTGAAG ATCCAGCCCA CAGATCCTGG GGACTCGGCC
VB8s2  -A-...---G ----A----- ---------- ---------- ---------- ------A---

         301         311
          .           .
VB8s1  ACGTACCTCT GT
VB8s2  ---------- --
                                    CDR1                          CDR2
```

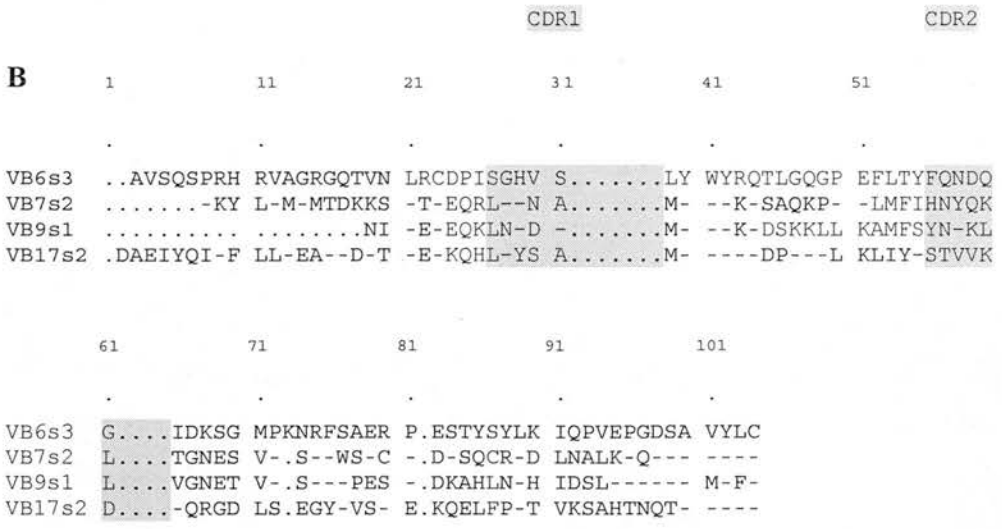
B

```
          1          11          21          31          41          51
          .           .           .           .           .           .
VB8s1  NLGVTQTPRH EVTEKGQAVT LSCELIKSHT A.......LY WYRQTSVRGL EFLAYFSNQA
VB8s2  ---------- ---------- ----P----A -........-F ---------- ---I--N---

          61          71          81          91         101
          .           .           .           .           .
VB8s1  P....IDESG MPKDRFSAKM P.NSSFSTLK IQPTDPGDSA TYLC
VB8s2  -....----- ---------- H.-A------ ---------- ----
```

E.5 Nucleotide (A) and amino acid (B) sequence of Vβ6s3, Vβ7s2, Vβ9s1 and Vβ17s2

A

```
          1          11          21          31          41          51
          .           .           .           .           .           .
VB6s3     ......GCAG TGTCCCAGTC CCCCAGGCAC AGGGTTGCAG GGAGGGGCCA GACTGTGAAT
VB7s2     .......... .......... .--A-AAT-- CTA--CATG- -A-T-ACGG- C-AGAAATC-
VB9s1     .......... .......... .......... .......... .......... ..AGAAC-TC
VB17s2    ...GAT--T- AAAT-T-CCA GATA--ATTT CT-C-C--T- A-GCT----- -GA-----CC

                                          CDR1
          61          71          81          91         101         111
          .           .           .           .           .           .
VB6s3     CTTAGGTGTG ACCCAATTTC TGGTCATGTT TCC....... .......... ....CTTTAT
VB7s2     T-G-CA---- -A-A-CG-CT G--A---AA- G-T....... .......... ....A-G--C
VB9s1     --AGAA---- -G-A--AGCT GAAC----A- AGT....... .......... ....A-G---
VB17s2    --GGA----A -A-AGCACCT G--CT-CAG- G--....... .......... ....A-G--C

                                                                  CDR2
         121         131         141         151         161         171
          .           .           .           .           .           .
VB6s3     TGGTACCGAC AGACCCTGGG GCAGGGCCCA GAGTTTTTGA CTTACTTTCA AAACGACCAA
VB7s2     -----TAA-- ---G-GCCCA -A--CCG--- ---C-CA--T TCAT-CACA- CT--C-GA--
VB9s1     -----TAAG- --GA-TCCAA -A-ATTG-TG A--GC-A--T T-AG--ACA- T--TA-G-T-
VB17s2    --------G- --GA--CA-- ------T-TG A--C-GA-TT A-----C-AC GGTT-TAA-G

         181         191         201         211         221         231
          .           .           .           .           .           .
VB6s3     GGA....... .....ATAGA CAAATCAGGG ATGCCTAAAA ACCGGTTTTC TGCTGAGAGG
VB7s2     CTC....... .....-CT-G A--TGAGA-T G-T--G...- GT--C--C-G GT----AT-C
VB9s1     CTC....... .....G-T-G A--TGA-AC- G----A...- GT--T--C-- AC-G---TCT
VB17s2    -AT....... .....--TC- G-G-GG--AC T-AT--...G -AG-C-ACAG --TCTCTC-A

         241         251         261         271         281         291
          .           .           .           .           .           .
VB6s3     CCT...GAGA GCACATACTC CTATCTGAAG ATCCAGCCAG TGGAGCCCGG GGACTCAGCA
VB7s2     --A...--C- --T-TC-G-G -CGC---G-C C-GA-CG-CC --A-A--ACA ------C--T
VB9s1     ---...--C- AAG-TC-T-T AA-C--TC-C ---G-CT-CC -------A-- C-----T--C
VB17s2    GAG...A--C AGGAGCTG-T TCC---C-CT G-GA--T--- CTC-CA--AA CC-AA----C

         301         311
          .           .
VB6s3     GTGTATCTCT GT
VB7s2     --C------- -C
VB9s1     A-----T--- --
VB17s2    -----C---- -C
```


CDR1
CDR2
B
1
11
21
31
41
51
VB6s3 ..AVSQSPRH RVAGRGQTVN LRCDPISGHV S.......LY WYRQTLGQGP EFLTYFQNDQ
VB7s2 ....KY L-M-MTDKKS -T-EQRL--N A.......M- --K-SAQKP- -LMFIHNYQK
VB9s1 ....NI -E-EQKLN-D -.......M- --K-DSKKLL KAMFSYN-KL
VB17s2 .DAEIYQI-F LL-EA--D-T -E-KQHL-YS A.......M- ----DP---L KLIY-STVVK
61
71
81
91
101
VB6s3 G....IDKSG MPKNRFSAER P.ESTYSYLK IQPVEPGDSA VYLC
VB7s2 L....TGNES V-.S--WS-C -.D-SQCR-D LNALK-Q--- ----
VB9s1 L....VGNET V-.S---PES -.DKAHLN-H IDSL------ M-F-
VB17s2 D....-QRGD LS.EGY-VS- E.KQELFP-T VKSAHTNQT- ----

Appendix F - Artefactual band from Vβ subfamily-specific snPCR

Alignment of the nucleotide sequences of (A) the snPCR artefactual band, (B) locus XM_595623 of bovine genome scaffold NW_932240.1/Bt9_WGA1710_2 and (C) the human nesprin-1 gene (GenBank accession number AF495910 - Appendix B). Identity is shown as dashes and gaps by dots. The potential binding sites of BCint at the 5' end and BCint (reverse complement) at the 3' end of the artefactual band are indicated by the shaded boxes

```
     1          11         21         31         41         51
     .          .          .          .          .          .
A    GGTCAGCTC. ..CACGTGGT CAACAAATTG AGACTGGTGG AACAAAAGTT TCAGCAGGTA
B    AC----T--T AA-------- ---------- ---------- ---------- ----------
C    AC----T--T AA-------- G--------- --G--AA--- -G-------- ------A---
                BCint
     61         71         81         91         101        111
     .          .          .          .          .          .
A    GATGAGTGGC TGAAAACAAC GGAGGAGAAA GTGAGTCTCA GAACGGCACG TCAGTCTAGC
B    ---------- ---------- ---------- ---------- ---------- ----------
C    -----A---- -C------G- A--------- --T----C-- -G--CAG--- --------A-

     121        131        141        151        161        171
     .          .          .          .          .          .
A    CGAGCCGCCA AGGAGATGCA GTTACATCAG ATGAAGAAGT GGCATGAAGA AATCACCGCA
B    ---------- ---------- ---------- ---------- ---------- ----------
C    A-G--AA--- -------A-- A--------- ---------- ----C----- -G-G--T---

     181        191        201        211        221        231
     .          .          .          .          .          .
A    TACAGAGATG AAGTTGAGGA AGTGGGAGCC AGAGCACAGG AGATCCTGGA CGAGAGCCAC
B    ---------- ---------- ---------- ---------- ---------- ----------
C    ---------- ---------- ---------T -----T---- ----A----- ----------

     241        251        261        271        281        291
     .          .          .          .          .          .
A    GTGAGCAGCC GAATGGGCTG CCAGGCCACG CAGCTGACGT CCAAATACCA AGCGCTGCTC
B    ---------- ---------- ---------A ---------- ---------- ----------
C    ----A----A -------T-- ---------C --------T- ---G------ G--C-----T
```

```
     301         311         321         331         341         351
      .           .           .           .           .           .
A    CTCCAAGTGC TGGAACAAAT AAAATTCCTG GAAGAGGAGA TCCAGAGTTT GGAGGAATCA
B    ---------- ---------- ---------- ---------- ---------- ----------
C    ---------- ---------- ---------- --G------- -T-------- ----------

     361         371         381         391         401         411
      .           .           .           .           .           .
A    GAGTTATCCT TAAGTTCCTA TTCTGATTGG TACAGCTCTA CTCATAAAAA CTTCAAGAAT
B    ---------- ---------- ---------- ---------- ---------- ----------
C    --A-C----C -C-------- ---------- --TG------ ---------- ----------

     421         431         441         451         461         471
      .           .           .           .           .           .
A    GTGGCTGCCA AAATTGATAA AGTAGATAAA GCGATGATGG GGAAAAAGAT GAAGACGCTG
B    ---------- ---------- ---------- ---------- ---------- ----------
C    ------A--- -G-----C-- --------C- -TA------- ----G--AT- -------T--

     481         491         501         511         521         531
      .           .           .           .           .           .
A    GAGGGTTTGC TGAAAGACAT GGAGAAAGGC CAGAGCTTGC TCAAATCAGC CCGAGAGAAA
B    ---------- ---------- ---------- ---------- ---------- ----------
C    ----T----- -C-------- ---------T --C--T---- -G-------- ---G------

     541         551         561         571         581         591
      .           .           .           .           .           .
A    GGCGAGAGGG CTCTTAAATA CTTGGAAGAT GCTGAGGCAG AGACGTTAAG AAAAGAGATC
B    ---------- ---------- -------C-- ---------- ---------- ----------
C    --A------- --G------- ------G--A -GC------- ---G------ ---G-----T

     601         611         621
      .           .           .
A    CACGACCACG TGGAGCGGTT G
B    ---------- ---------- -
C    --T--T---A ------A--- -
        <-------------------
        BCint (reverse complement)
```

Appendix G - Sequence nucleotide identity between Vβ sequences in the genome

Vβ genes sharing ≥98% nucleotide sequence homology are, by convention, assumed to represent alleles of the same gene (Arden *et al.*, 1995a, b). After editing to remove the leader sequence, intron and the recombination signal sequence (RSS), all Vβ genes identified in the genome were compared for percentage nucleotide sequence identity to the other members of their subfamily (data not shown) and genes sharing ≥98% nucleotide sequence identity noted. Unless genes sharing ≥98% were (i) serially located on the same contig or (ii) adjacent to different Vβ genes at equivalent positions in their respective contigs, it was not possible to classify whether the sequences represent distinct gene loci or allelic variants. The results of this analysis are summarized in Table G.1 below. The status of fifteen gene sequences (VB1ag, ad, ae, af, VB2e, VB7c, VB10j, s, k, l, t VB13ah, ai, al and VB17a) were considered to be unclassified and excluded from further analysis.

Vβ subfamily	Genes	Nucleotide identity (%)	Comments
VB1	c/ag*	100	1ag removed
	d/z/ad	99-100	d/z not alleles - adjacent to different Vβ genes 1ad removed
	e/y	99	Not alleles - adjacent to different Vβ genes
	m/ae*	100	1ae removed
	p/aa	100	Not alleles - adjacent to different Vβ genes
	v/w	100	Not alleles - serially located on same contig
	x/ab	99	Not alleles - adjacent to different Vβ genes
	ac/af	100	1af removed
VB2	a/e*	100	2e removed
VB7	a/c	99.7	7c removed
VB10	e/j	98.6	10j removed
	i/s	99.0	10s removed

Table G.1 Combinations of Vβ genes identified in the bovine genome that share ≥98% nucleotide sequence homology. (*) identified Vβ genes of which the genomic sequence was incomplete.

Vβ subfamily	Genes	Nucleotide identity (%)	Comments
VB10	k*/r	99.4	10k removed
	k*/l*	98.9	10k removed
	l*/q*	100	10l removed
	l*/r	99.4	10l removed
	m/t*	100	10t removed
VB13	d/ah*	100	13ah removed
	e/ai/ag	98.3-99.0	e/ag not alleles - adjacent to different Vβ genes 13ai removed
	j/ab*	100	Not alleles - adjacent to different Vβ genes
	s/al*	100	13al removed
	z*/af	100	Not alleles - serially located on same contig
	g/aj	99.7	Not alleles - adjacent to different Vβ genes
	ac/ad	99.0	Not alleles - serially located on same contig
VB17	a*/b	98.6	17a removed
	a*/c	100	17a removed
	a*/d*	98.6	17a removed

Table G.1 (continued) Combinations of Vβ genes identified in the bovine genome that share ≥98% nucleotide sequence homology. (*) identified Vβ genes of which the genomic sequence was incomplete.

Appendix H - Analysis of the functionality of Vβ gene sequences identified in the genome

Sequenced data for the entire Vβ genes identified from the genome from initiation codon of the leader (L) exon to the end of the RSS was analysed to enable prediction of the functional status of the genes according to the criteria described in section 5.2.3. Tables H.1 and H.2 summarize the results of this analysis.

Table H.1 (overleaf) Genomic Vβ sequence analysis for assessment of functionality (excluding the portion of the V exon encoding the mature Vβ polypeptide).

(i) **Bovine gene name. (*) denotes VB gene with incomplete sequence data.**

(ii) **Translation of the L exon. (*) denotes stop codon. Some of the L exons had insertions/deletions that would cause a frameshift in the V exon after splicing and consequent introduction of premature stop codons.**

(iii) **Splice site sequence. The splice sites flanking the intron were assessed for the presence of the conventional GT/AG splice site motif. The 5'flanking splice site is shown above the 3'flanking splice site**

(iv) **Translation of the portion of the V exon encoding leader sequence. This was predicted following alignment of the genomic Vβ sequences using the IMGT/V-QUEST software programme (http://imgt.cines.fr - Giudicelli *et al.*, 2004)**

(v) **Intron length.**

(vi) **Length of sequence in the germline arrangement between the codon for the conserved cysteine at position 104 in the Vβ gene and the start of the RSS.**

(vii-ix) **Sequence of the components of the RSS. Italicised, underlined text with an asterisk indicates spacer sequences that vary by +/- 1bp from the normal 23bp.**

Shaded text indicates components of individual VB gene sequences which were considered to render them non-functional. The full analysis of the portion of the V exon encoding the mature Vβ polypeptide has not be included due to its volume - details of any V exon defects leading to non-functional VB sequences have been summarised in Table H.2. The name of the all VB genes considered to be non-functional (including those considered non-functional due to defects in the V exon as detailed in Table H.2) are indicated by shading. The information contained in columns (iv), (v) and (vi) are not directly related to functional status but were included in the general analysis of VB gene sequence data and so are presented here.

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Bovine Gene	Leader sequence from L exon	Splice sites	Leader sequence from V exon	Intron length	CDR3 Length (bp)	Heptamer	Spacer	Nonamer
1a	MGSRLLCCVTLCLLGA	GTGAG AACAG	GLV	131	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ACAAAAACA
1b	MGSRLLCCVTLCHLGA	GTGAG AACTG	GLV	132	14	CACAGCC	CTGCATGACCAGGGGCCTCTGGG	TCAAAAACA
1c	MGSRLLCFVTLCLLGA	GTGAG AAAAG	GLV	133	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ATAAAAACT
1d	MGSRLLCCVTLCLLGA	GTGAG AACAG	GLV	131	14	CACAGCC	CTGCATGATCAGATGCCTCTGGT	ACAAAAACA
1e	MGSRLLCCVTLCLLGA	GTGAG AACAG	GLV	134	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ACAAAAACA
1f	MGSRLLCCVTSSFWEQ	GTGAG AGCAG	-	129	14	CACAGCC	CTACATGATCAGGTGCCTCTGGA	ACAAAAACA
1g	MGCSPVYCAALCLLPA	GTGAG CACAG	GLV	125	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ACAAAAACA
1h	MDCSRVYCVALWLLAA	GTGGG CACAG	GLV	115	14	CACAGCC	CTGCATGATCAGCTGCCTCTGGG	ACAAAAACA
1i	MGSRLLCCVTLCLLGA	GTGAG AACAG	GLV	133	14	CACAGTG	CTGCATGATAAGGAGCCTCTGGG	ACAAAAACA

Table H.1. Genomic Vβ sequence analysis for assessment of functionality (1 of 11 pages)

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
1j	MGSRLLCCVTLCLLGA	GTGAG AACAG	GLV	132	14	CACAGCC	CTGCATGATCAGGTGCCTGTGGG	ACAAAAACA
1k	MGSRLLCCVTLCLLGA	GTGAG AACAG	GLV	134	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ACAAAAACA
1l	MGSRLLCCVTSSFWE*	GTGAG AAGAG	-	133	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGT	ACAGAAACA
1m	MGCSPVCCVALCLLAA	GTGGG CACAG	GLV	133	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ACAAAAACA
1n	MGSRLL*CVTLCLLGA	GTGAG AACAG	-	131	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	AGAAAAACA
1o	MGSRLLCCVTLCLLGA	GTGAG AGCAG	GLV	129	14	CACAGCC	CTGCATGATCAGATGCCTCTGGA	ACAAAAACA
1p	MGSRLLCCVTLCLLGA	GTGAG AACAG	GLV	117	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	GCAAAAACT
1q	MGSRLLCCVTLCLLGA	GTGAG CACAG	GLV	131	14	CACAAGC	*CCTGCATGATCAGGTGCCTCTGGG**	ACACAAACT
1r	MGSRFLCCVTLCLLGA	GTGAG AACAG	GLV	129	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ACAAAAACA
1s	MGSRPLCCVTLCLLGA	GTGAG AACAG	GLV	117	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ACAAAAACA
1t	MGSRLLCCVTSSFCEQ	GTGAG AACAG	-	131	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGT	ACAGAAACA
1u	MGCSPVCCVALCLLAA	GTTGG CACAG	GLV	131	14	CACAGCC	CTGCATGATCAGTTGCCTCTGGG	ACAAAAACA

Table H.1. Genomic Vβ sequence analysis for assessment of functionality (2 of 11 pages)

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
1v	MGSRLL*CVTLCLLGA	GTGAG AACAG	-	133	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ACAAAAACA
1w	MGSRLLCCVTLCLLGA	GTGAG AACAG	GLV	135	14	CACAGCC	*CCTGCATGATCAGGTGCCTCTGGG**	ACAAAAACA
1x	MGCSPVCCVALCLLAA	GTTGG CACAG	GLV	133	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ACAAAAACA
1y	MGSRLLCCVTLCLLGA	GTGAG AACAG	GLV	131	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ACAAAAACA
1z	MGSRLLCCVTLCLLGA	GTGAG AACAG	GLV	129	14	CACAGCC	CTGCATGATCAGATGCCTCTGGT	ACAAAAACA
1aa	MGSRLLCCVTLCLLGA	GTGAG AACAG	GLV	115	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	GCAAAAACT
1ab	MGCSPVCCVALCLLAA	GTGGG CACAG	GLV	132	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ACAAAAACA
1ac	MGPRLLCCVTLCLLGA	GTGAG AACGG	GLV	132	14	CACAGCC	CTGCATGATCAGGGGCCTCTGAG	ACAAAAACA
2a	MLVLLLLLGP	GTCAG AGCAG	SSGL	499	11	CACAGCG	CTGGGCAGAGATCAGAGACCCAG	ACAAGAATC
2b	MLVLLLLLGP	GTCAG AGCAG	SSGL	479	11	CACAGCG	CCGGGCAGAGATCAGAGACCCAG	ACAAGAATC
2c	MLVLLLLLGP	GTCAG GACAG	-	378	11	CACAGCG	CTGGGCAGAGATCAGAGACCCAG	ACAAGAATC
2d*	No sequence	AGCAG	-	-	11	CACAGCG	CTGGGCAGAGATCAGAGACCCAG	ACAAGAATC

Table H.1. Genomic Vβ sequence analysis for assessment of functionality (3 of 11 pages)

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
2f	MLVLLLLLGP	GTCAG AGCAG	SSGL	482	11	CACAGCG	CTGGGCCGAGATCAGAGACCCAG	ACAAGAATC
3a	MGIRLLYAFCFLGV	GTGAG CACAG	GLV	335	14	CACAGCG	CAGCACGGCCACATCCTCTCTGC	ACAAAAAGG
4a	MLTLLLLLLGL	GTATG TCTAG	GSVF	261	11	CACAGTG	CTGGGCACAGTTCAAGGGTCTCA	GCAAGAACC
4b	MLTLLLLLLGL	GTATG TCTAG	GSVF	261	11	CACAGTG	CTGGGCACAGATCAAGGGTCTCA	GCAAGAACC
4c	MLTLLLLLLGL	GTACG TCTAG	GSVF	259	11	CACAGTG	CTGGGCACAGATCAAGGGTCTCA	GCAAGAACC
4d	MLTFLLLLLEL	GTATG TCTAG	GSVF	304	11	CACAGCG	CTGGGCACAGATCAAGGGTCTCA	GGAAGAACC
5a	MGCSPVCCVALCLVAA	GTGAG CACAC	GLV	114	14	CAGATCC	ATGAGAGCCTGGGGACTGCAGTGGG	AGACAAAAA
5b	MGSRLLCCVTLCLLGA	GTGAG AACAG	GLV	131	14	CACAGCC	CTGCATGATCAGGTGCCTCTGGG	ACAAAAACA
5c	MGSRLLPWVMLYLLRA	GTAAG CACAG	GPV	125	13	CACAGCT	GTGCTGAGTCACCCTCGCTCTGC	ACACAAACC
5d	MGSRLLCCVTSSFWEQ	GCGAG AACAG	-	134	14	CACAGCC	CTGCATGATTAGGTGCCTCTGGA	ACAAAAACA
6a	MGSRLLCWATLCLLGV	GTGAG CACAG	GHT	154	14	CACAGCG	CAGCACAGTCGCCTCCTTCCTGT	TCACAAACC

Table H.1. Genomic Vβ sequence analysis for assessment of functionality (4 of 11 pages)

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
6b	MGSRLLCWVTLCVLGV	GTGAG CACAG	GHT	292	14	CACAGCG	CGGCACAGTCACCTCCTTCCTGT	TCACAAACC
7a	MGFRQLCCVVLCLLGI	GTGAG CCCAG	AVCA	119	14	CACAGCC	CTGCAGAGGCAATTCCTTCCCGT	GCACAAACC
7b	MGFRQLCCVVLCLLGI	GTGAG CCCAG	AVCA	119	14	CACAGCC	CTGCAGAGGCAATTCCTTCCTGT	GCACAAACC
8a	MGTRALCSMVLCLLGV	GTGAG CACAG	GTT	91	14	CACAGCG	*CTGCAGAGTTGCCCCGTCCCGA**	GCAGAAAAC
8b	MATKLLCCVALCLLGV	GTAAG CACAG	GPT	96	14	CACAGCG	*CTACAGAGTTGCCCCCTCCCGA**	GCAGAAAAC
9a	MDSRLLCCVALGLLGV	GTGAG CTCAG	GTW	117	14	CACAGCC	CTGCAGAGTCACTGCCTGCTCGT	GCACAAAGC
10a	MCFSLLCCSAFFSSGEQ	GTAAA CACAG	-	104	14	CACAGTG	CTGCACGTCAGCTCTCCTTAGAG	CACAAACTC
10b	MCFGLLCCVALFIWGA	GTGAG TACAG	GSM	105	14	CACAGTG	CTGCGCGTCAGCTCTTCTTAGAG	CACAAACTC
10c*	MCLSLLCCVALFLWGA	GTAAG CACAG	CSM	112	-	CACAGTG	CTGCACGTCAGCTCTTCTTACAG	CACAAACTC
10d	VCFSLLCCAAFFSSKE	GTATA TACAG	-	86	14	CATAGTG	CTGCGCGTCAGCTTTTCTTAGAG	CACAAACTC
10e	MSLSLLCWVALFLWGA	GTGAG TACAG	GSM	86	14	CACAGTG	CTGCACGTCAGCTCTTCTTAGAA	CACAAACTC
10f	MCLGLLCCMALLFWGA	GTGAG CACAG	GSM	104	14	CATAGTG	CTGCGCGTCAGCTCTTCTTAGAG	CACAAACTC

Table H.1. Genomic Vβ sequence analysis for assessment of functionality (5 of 11 pages)

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
10g	MCLGLLCCMALLFWGA	GTAAG CACAG	GSM	112	14	CACAGTG	CTGCGCGTCAGCTCTTCTTAGAG	CACAAACTC
10h	MCFSLLFCAAFFFFWGA	GTAAG TACAG	GSM	102	14	CACAGTG	CTGCGCGTCAGCTCTTCTTAGAA	CACAAACTC
10i	MCLGLLCCVALFIWGA	GTGAG TACAG	GSM	86	14	CACAGTG	CTGCGCGTCAGCTCTTCTTAGAG	CACAAACTC
10m	MCLGLLCCVALFIWGA	GTGAG TACAG	GSM	104	14	CACAGTG	CTACGCGTCAGCTCTTCTTAGAG	CACAAACTC
10n	MCFSLLCCAAFCFLLGSR	GTGAG TACAG	-	102	14	CACAGTG	CTGCGCGTCAGCTCTCCTTAGAG	CACAAACTC
10o	MCFSLLCCAAFFFFWGA	GTAAA TACAG	SSM	112	14	CACAGCG	CTGCACGTCAGCT---CTTAGAG	CACAAACTC
10p	MSLSLLCWVALFLWGA	GTGAG TACAG	GSM	100	14	CACAGTG	CTGCACATCAGCTCTTCTTAGAG	CACAAACTC
10q*	MSLSLLCWVALFLWGA	GTGAG TACAG	-	100	14	CACAGTG	CTGCACATCAGCTCTTCTTAGAG	CACAAACTC
10r	MCFSLLCCAAFCFLLGSR	GTGAG TACAG	-	104	14	CACAGTG	CTGCACGTCAGCTCTCCTTAGAG	CACAAACTC
10u*	No sequence		-	-	14	CACAGTG	CTGCACGTCAGCTCTTGTTCGAA	CACAAACTC
11a	MAVRLLCSVALYLLGA	GTGAG CTCAG	GFS	127	14	CACAGCA	CGACACAGGCCCCCACAGACTGC	ACAAAAAGG
12a	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	73	14	CATGGCA	*CTGCATGGCTACCTGCTTCTGT**	GCAAAAAGA

Table H.1. Genomic Vβ sequence analysis for assessment of functionality (6 of 11 pages)

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
13a	ISPCLLCCVVFCLLQA	GTGAG CACAG	-	88	14	CACAGCG	*CTGCAGGGTCACCTGCTCTCTGTT**	GAAAAAGAC
13b	MSLGLLGCVVFCFLQA	GTGAG CACAA	GAV	87	14	CACGGCA	CTGCAGGGTCACCTGCTCTCTGT	GCAAAAGAC
13c	MSPRLLGCVVFCLLQA	GTGAG CACAA	GVV	86	14	CACGGCG	CTGCATGGCCACCTCCTCTCTGT	GCAAAAAAA
13d	MSPCFLGCVVFYLLPA	GTGAG CACAG	GAV	86	13	CACGGTG	CTGCATGGCCACCTCCTCTCTGT	GCAAAAAGA
13e	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	13	CACAGCG	CTGCAGGGTCACCTCCTCTCTGT	GCAAAAAGG
13f	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	86	13	CACGGCG	CTGCACGGCAACCTCCTCTCTGT	GCAAAAAGA
13g	MSPCFLGCVVFCLLQA	GTGAG CACAG	GAV	86	14	CACGGTG	CTGCACGGCCACCTCCTTTCTAT	GCAAAAAAA
13h	MNPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	14	CACAGCG	CTGCAGGGTCACCTCCTCTCTGT	GCAAAAAGA
13i	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	86	14	CACGGTG	CTGCACGGCCACCTCCTTTCTGT	GCAAAAAAG
13j	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	14	CACAGCG	CTGCGTGGCCACCTCCTCTCTGT	GCAAAAAGA
13k	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	13	CACAGCG	CTGCAGGGTCACCTCCTCTCTGT	GCAAAAAGA
13l	MSPCLLGCVVFCLLQT	GTGAG CACAG	GAV	87	14	CACGGTG	CTGCACGGCCACCTGCTTTCTAT	GCAAAAAAG

Table H.1. Genomic Vβ sequence analysis for assessment of functionality (7 of 11 pages)

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
13m	VSPCLLGCVVFCLLQA	GTGAG CACAG	-	87	14	CACAGCG	CTGCAGGGTCACCTGCTCTCTGT	TGAAAAAGA
13n	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	13	TACAGCG	CTGCACGGCCACCTCCTCTCTGT	GCAAAAACA
13o	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	86	14	CACAGTG	CTGCATGGCCACCTCCTTTCTGT	GCAAAAAAG
13p	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	14	CATGGTG	CTGCACGGCCACCTCTTTTCTGT	GCAAAAAAG
13q	MSPCLLGCVVFCLLQA	GCGAG CACAG	GTV	86	13	TGCGGCG	CTGCTTGGACACCTCCTCTCTGT	GCAATAAGA
13r	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	14	CACAGCG	CTGCAGGGTCACCTCCTCTCTGT	GCAAAAAGG
13s	MSPCFLGCVVFCLLQA	GTGAG CACAG	GVV	87	14	CATGGTG	CTGCATGGCCACCTCTTCTCTGT	GCAAAAATA
13t	MSPCLLGCVVFCLLQA	GTGAG CACAA	GAV	87	14	CACGGCA	CTGCACGGCCACCTCCTCTCTGT	GCAGAAAGA
13u	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	14	CACAGCG	CTGCAGGGTCACCTGCTCTCTGT	TGAAAAAGA
13v	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	86	13	CACGGCG	CTGCACGGCCACCTCCTGTCTGT	GCAAAAAGT
13w	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	13	CACAGCG	CTGCATGGCCACCTCCTCTCTGT	GCAAAAAGA
13x	MSPCHLGCVVFCLLQA	GTGAG CACAG	GAV	87	14	CACAGTG	CTGCATGGCCACCTCCTTTCTGT	GCAAAAAAG

Table H.1. Genomic Vβ sequence analysis for assessment of functionality (8 of 11 pages)

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
13y	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAI	87	14	CACGGCG	CTGCATGGCCACCTGCTCTCTGT	GCAAAAAGA
13z*	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	86	No sequence			
13aa	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	14	CATGGCG	CTGCACGGCCACCTGCTCTCTGT	GCAAAAAGA
13ab*	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	No sequence			
13ac	MSPCLLSCVVFCLLQA	GCGAG CACAG	GAV	87	14	CATGGCG	CTGCACGGCCACCTGCTCTCTGT	GCAAAAAGA
13ad	MSPCLLGCVVFCLLQA	GCGAG CACAG	GAV	87	14	CACGGAG	CTGCATGGCCACCTGATCTCTGT	GCAAAAAGA
13ae	MSPCLLGCVVFCLLQA	GTGAG CACAG	GPV	87	14	CACGGTG	CTGCACAGCCACTCCTTTCTAT*	GCAAAAAGC
13af	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	86	13	CACGGTG	CTGCACGGCCACCTCCTCTCTGT	GCAAAAAGA
13ag	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	14	CACAGCG	CTGCAGGGTCACCTCCTCTCTGT	GCAAAAAGG
13aj	MSPCFLGCVVFCLLQA	GTGAG CACAG	GAV	87	14	CACGGTG	CTGCACGGCCACCTCCTTTCTAT	GCAAAAAAA
13ak	MSPCLLGCVVFCLLQA	GTGAG CACAG	GAV	87	14	CACGGCG	CTGCATGGCCACCTCCTCTCTGT	GCAAAAAGA
13am	MSPCLLGCLVFCLLQA	GTGAG CACAG	GTV	87	14	CACGGCG	CTGCATGGCCACCTGCTCTCTGT	GCAAAAAGA

Table H.1. Genomic Vβ sequence analysis for assessment of functionality (9 of 11 pages)

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
14a	MGPWLLGCFLFYLLGA	GTGAG CACAG	GPL	127	14	CACAGTG	TGACACAGCCAGCTGCTCTCTCC	ACAAAAAGG
15a	MLLCYVIFCLLGT	GTGAG CACAG	GSM	87	14	CACAGTG	CTTCCTGGCCACCTGCTCTCTAC	ACAGAAAGA
16a	MLSSLLRLASLCLLGA	GTGAG CACAG	RYI	82	14	CACAGTG	CTTCAGAGTTCTGTCCTTTCTGT	GCAAAACCA
17b	MGNQVI*CVALCLFGA	GTGAG CATAG	-	133	14	CACAGTG	ACTCAAAGCCACCTCCTCTCTGC	ACACAAATG
17c	MGNQVICCVALCLFGA	GTGAG CATAG	GTV	133	14	CACAGTG	ACTCAAAGCCACCTCCTCTCTGC	ACACAAATG
17d*	MGNQVICCVALCLFGA	GTGAG CATAG	GTV	133	14	CACAGTG	ACTCAAAGCCACCTCCTCTCTGC	ACACAAATG
17s2a	MGSWALYLALCLLGA	GTGAG TACAG	GPM	133	11	CACAGTG	GAGCACGGCCACCTCTCACCTGT	GCACAAACC
18a	MGTSLLS*GVIWLLEA	GTGAG CTGAG	-	284	14	CATATCA	ATGCAG--TCATATCCTTTCACT	GCACAAACA
18b	MGTSLLS*GVICSWWQ	GTGAG CTGAG	-	287	14	CATATCA	ATGCAG--TCATGTCCTTCCACT	GCACAAACA
18c*	No sequence		-	-		CATATCA	ACGCAG--TCATGTCCTTCCACT	GCACAAACA
20a	MPCCLLALLLGTFL	GTGGG CGCAG	GVR	326	11	CACACTG	CGCTGGGTGGGGCAGACATCTGT	GCAGAAACC
21a	MGSSLLFWVTLCLLGA	GTGAG GGCAG	ELT	110	14	CATAGCC	TCACAGAGACCCTTCCTTTCTGT	GCACAAATT
23a	KPSLLLPELTRVLGFLAV CPAVS*EQ	AGGAG TCCAG	-	84	13	CACAGCC	CTGCAAAGTCCCTGGCTTTCTGT	ACTCAAACC

Table H.1. Genomic Vβ sequence analysis for assessment of functionality (10 of 11 pages)

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
24a	MRPRLLLCVALYLLGA	GTAAG TGCAG	GHV	121	14	CACAGAG	CTGAAGCCTTCCCTGCCCTCTGT	TCATAAACC
25a	TNPTLTCFTVLCLLGA	GGAAG CACAG	GFL	107	14	CACAGTG	AGAAATACCGGCTAAGCTTAATG	CACACTCAG
28a	MCNRLLCCVVLLLLRV	GTGAG CACAG	GLK	136	14	CACAGCG	TGGCACGGCTGCTTCCTCTCTGC	ACAAAAAGA
Xsa	MAPFPILWGWVLLAAVS T	GTCAG TTCAG	DAA	274	14	CACAGTG	AGACACCCTAACTGGGAAAATGA	ATAAAAACC
Untitled 1*	No sequence	CACAG	-	-		CACAGCG	TGGCATAGCCAGCTGCTCTCTGC	ACAAAAAGG
Untitled 2	MVIRLLCDFVSPVGK	GTAAG CACAG	-	85		TATAGTG	CTGCATGGCCACTTCCTCTCTGC	ACAGAAGGG

Table H.1. Genomic Vβ sequence analysis for assessment of functionality (11 of 11 pages)

Table H.2 Summary of non-functional VB genes. This table summarises the justification for the classification of VB gene sequences as either pseudogenes, ORF pseudogenes or orphan pseudogenes.

Vβsubfamily	Gene	Functionality	Reason
1	b	Pseudogene	Frameshift in V exon, defect in acceptor splice
	f	Pseudogene	Frameshift in L exon, Frameshift in V exon
	l	Pseudogene	Frameshift in L exon, Frameshift in V exon
	n	Pseudogene	Stop codon in L exon
	t	Pseudogene	Frameshift in L exon, Frameshift in V exon
	v	Pseudogene	Stop codon in L exon
	x	Pseudogene	Frameshift in V exon
	ab	Pseudogene	Frameshift in V exon
	ac	ORF	Defect in acceptor splice
2	c	ORF	Extra codons in V exon
4	d	Pseudogene	Frameshift in V exon
5	a	Pseudogene	Frameshift in V exon, defect in acceptor splice, defect in RSS
	b	Pseudogene	Stop codon in Vexon
	d	Pseudogene	Frameshift in L exon, Frameshift in V exon, defect in acceptor splice
10	a	Pseudogene	Frameshift in L exon
	d	Pseudogene	Frameshift in L exon, Frameshift in V exon, defect in RSS
	f	Pseudogene	Frameshift in V exon, defect in RSS
	h	Pseudogene	Stop codon in V exon
	n	Pseudogene	Frameshift in L exon, Frameshift in V exon
	o	ORF	Defect in RSS
	r	Pseudogene	Frameshift in L
12	a	Pseudogene	Stop codons in V exon, defect in RSS

Table H.2 (continued) Summary of non-functional VB genes.

13	a	Pseudogene	Loss of initiation codon
	b	Pseudogene	Frameshift in V exon, defect in acceptor splice
	c	Pseudogene	Frameshift in V exon, defect in acceptor splice, defect in RSS
	d	ORF	Loss of conserved TRP@41
	f	ORF	Loss of conserved Cys@104
	g	ORF	Defect in RSS
	l	Orphan	Located on chromosome 5
	m	Pseudogene	Loss of initiation codon, loss of conserved Cys@104
	n	ORF	Defect in RSS
	o	Pseudogene	Frameshift in V exon
	p	ORF	Defect in RSS
	q	Pseudogene	Frameshift in V exon, defect in acceptor splice, defect in RSS
	s	ORF	Defect in RSS
	t	Pseudogene	Stop codon in V, frameshift in V exon, defect in acceptor splice
	u	Pseudogene	Stop codon in V exon
	z	Pseudogene	Stop codon in V exon
	aa	ORF	Defect in RSS
	ac	ORF	Defect in acceptor splice, defect in RSS
	ad	ORF	Defect in acceptor splice
	af	Pseudogene	Stop codon in V exon
	aj	ORF	Defect in RSS
17	b	Pseudogene	Stop codon in L exon
18	a	Pseudogene	Frameshift in L exon, Frameshift in V exon, defect in RSS
	b	Pseudogene	Frameshift in L exon, Frameshift in V exon, defect in RSS
	c	Pseudogene	Frameshift in V exon, defect in RSS
21	a	Pseudogene	Frameshift in V exon, defect in RSS
23	a	Pseudogene	Loss of intiation codon, Frameshift in L exon, Frameshift in V exon, defect in acceptor splice
25	a	Pseudogene	Loss of initiation codon, defect in acceptor splice
Untitled	1	Pseudogene	Frameshift in V exon
	2	Pseudogene	Frameshift in L exon, Frameshift in V exon, defect in RSS

Appendix I - Nucleotide sequence comparison of human and bovine Jβ gene orthologues

The percentage nucleotide sequence identity of the genomic sequences of each bovine Jβ gene and the human Jβ gene with which it displays highest identity (i.e. its orthologue) is shown in Table I.1.

Bovine Jβ gene	Human orthologue	Nucleotide sequence identity (%)
Jβ1s1	Jβ1s1	89.6
Jβ1s2	Jβ1s2	72.1
Jβ1s3	Jβ1s4	72.5
Jβ1s4	Jβ1s5	70.0
Jβ1s5	Jβ1s6	83.0
Jβ2s1	Jβ2s1	82.0
Jβ2s2	Jβ2s2	74.5
Jβ2s3	Jβ2s4	85.7
Jβ2s4	Jβ2s5	85.4
Jβ2s5	Jβ2s7	70.2
Jβ3s1	Jβ2s1	78.0
Jβ3s2	Jβ2s2	74.5
Jβ3s3	Jβ2s2	67.3
Jβ3s4	Jβ2s3	85.7
Jβ3s5	Jβ2s5	84.8
Jβ3s6	Jβ2s6	79.2
Jβ3s7	Jβ2s7	80.9

Table I.1 Nucleotide sequence identity between bovine Jβ genes and their human orthologues.

Appendix J - Genomic bovine Cβ gene data

J.1 Location of the 3 bovine Cβ genes on scaffold NW_931380.1/Bt4_WGA850_2

Cβ	Exon	Length	Location on NW_931380.1/Bt4_WGA850_2
1	1	390	127,946 – 128,335
	2	18	127,261 – 127,278
	3	107	126,998 – 127, 104
	4	18	126,655 – 126,672
2	1	390	118,118 – 118,507
	2	*Region of undetermined sequence*	
	3	107	117,643 - 117,749
	4	18	117,299 – 117,316
3	1	390	109,509 – 109,898
	2	*Region of undetermined sequence*	
	3	107	108,748 – 108,854
	4	18	108,457 – 108,474

J.2 Alignment of the A) nucleotide and B) amino acid sequences of the genomic sequences of the 3 bovine Cβ genes.

Identity is shown as dashes and gaps (due to absence of sequence for exon 2 of bovine Cβ2 and Cβ3) by dots.

A

```
     1          11         21         31         41         51
     .          .          .          .          .          .
Cβ1 ATGATCTGAG CCGGGTCCAC CCGCCCAAGG TGGCTGTGTT CGAACCCTCG GAAGCAGAGA
Cβ2 ---------- ---------- ---------- ---------- ---------- ----------
Cβ3 ---------- ---C------ ---------- ---------- T--------- ----------

     61         71         81         91         101        111
     .          .          .          .          .          .
Cβ1 TCTCCCGGAC CCAGAAGGCC ACGCTCGTGT GCCTGGCCAC AGGCTTCTAC CCCGACCACG
Cβ2 ---------- ---------- ---------- ---------- ---------- ----------
Cβ3 ---------- ---------- ---------- G--------- --T------- ----------
```

```
    121         131         141         151         161         171
    .           .           .           .           .           .
Cβ1 TGGAGCTGAC CTGGTGGGTG AACAGGAAGC AGGTCACAAC TGGGGTCAGC ACGGACCCTG
Cβ2 ---------- ---------- ---------- ---------- ---------- ----------
Cβ3 ---------- ---------- ---------- -------G-- ---------- ----------

    181         191         201         211         221         231
    .           .           .           .           .           .
Cβ1 AGCCCTATAA GGAGGACCCT GCACGGGATG ACTCCAGATA CTGTCTGAGC AGCCGGCTGA
Cβ2 ---------- ---------- ---------- ---------- ---------- ----------
Cβ3 ---------- ---------C --GA------ ---------- ---------- ----------

    241         251         261         271         281         291
    .           .           .           .           .           .
Cβ1 GGGTGACCGC GGCCTTCTGG CACAACCCCC GCAACCACTT CCGATGCCAA GTCCAGTTCC
Cβ2 ---------- ---------- ---------- ---------- ---------- ----------
Cβ3 ---------- ---------- ---------- ---------- ---------- ----------

    301         311         321         331         341         351
    .           .           .           .           .           .
Cβ1 ACGGGCTCAC AGACCAGGAC CAGTGGGAGG AGCAGGACAG GGCCAAGCCC GTCACCCAGA
Cβ2 ---------- ---------- ---------- ---------- ---------A ----------
Cβ3 ---------- ---------- ---------- -----A---- -A-------- A---------

    361         371         381         391         401         411
    .           .           .           .           .           .
Cβ1 ACATCAGCGC CGAGGCCTGG GGCAGAGCAG ACTGTGGCGT CACCTCCGCA TCTTATCAGC
Cβ2 ---------- ---------- ---------- .......... .........-- ----------
Cβ3 ---------- ---------- ---------- .......... .........G ----------

    421         431         441         451         461         471
    .           .           .           .           .           .
Cβ1 AAGGCGTCCT GTCTGCCACC CTCCTCTATG AGATCCTGCT GGGGAAGGCC ACCCTGTACG
Cβ2 ---------- ---------- ---------- ---------- ---------- ----------
Cβ3 ----G----- ---------- -G-------- ---------- ---------- ----------

    481         491         501         511         521         531
    .           .           .           .           .           .
Cβ1 CTGTGCTGGT CAGCGCCCTG GTGCTGATGG CCATGGTCAA GAGAAAAGAG TCCTGA
Cβ2 ---------- ---------- -----A---- ---------- ---------- ------
Cβ3 ---------- ---------- ---------- -------T-- --A---G--T ------
```

B

```
    1           11          21          31          41          51
    .           .           .           .           .           .
Cβ1 DDLSRVHPPK VAVFEPSEAE ISRTQKATLV CLATGFYPDH VELTWWVNRK QVTTGVSTDP
Cβ2 ---------- ---------- ---------- ---------- ---------- ----------
Cβ3 ---------- ---------- ---------- ---------- ---------- ----------

    61          71          81          91          101         111
    .           .           .           .           .           .
Cβ1 EPYKEDPARD DSRYCLSSRL RVTAAFWHNP RNHFRCQVQF HGLTDQDQWE EQDRAKPVTQ
Cβ2 ---------- ---------- ---------- ---------- ---------- ----------
Cβ3 ---------- ---------- ---------- ---------- ---------- --N-T--I--

    121         131         141         151         161         171
    .           .           .           .           .           .
Cβ1 NISAEAWGRA DCGVTSASYQ QGVLSATLLY EILLGKATLY AVLVSALVLM AMVKRKES*
Cβ2 ---------- .......--- ---------- ---------- ---------- ---------
Cβ3 ---------- .......--- ---------- ---------- ---------- -----K-D-
```

J.3 Percentage nucleotide sequence identity comparisons of bovine and human Cβ gene exons

A) Cβ gene exon 1 (upper right) and exon 2 (lower left). * No sequence

	human Cβ1	**human Cβ2**	**bovine Cβ1**	**bovine Cβ2**	**bovine Cβ3**
human Cβ1		98.7	82.3	81.7	80.7
human Cβ2	88.9		81.2	81.2	80.5
bovine Cβ1	83.0	94.4		99.7	97.2
bovine Cβ2*	-	-	-		96.9
bovine Cβ3*	-	-	-	-	

B) Cβ gene exon 3 (upper right) and exon 4 (lower left)

	human Cβ1	**human Cβ2**	**bovine Cβ1**	**bovine Cβ2**	**bovine Cβ3**
human Cβ1		92.5	90.6	89.7	91.5
human Cβ2	90.4		88.8	87.8	88.8
bovine Cβ1	80.9	85.7		99.1	97.2
bovine Cβ2	90.4	85.7	100		96.3
bovine Cβ3	85.7	90.4	85.7	85.7	

Appendix K - Cytotoxicity results

Results of cytotoxicity levels ≥10% are considered to represent significant cytotoxic activity and are indicated by shading of the text. ND indicates cytotoxic assays for which no data was presented.

A) 592 clones

Clone Number	MHC-matched TpM	MHC-mismatched TpM	Tp2.1 presented by A10^{+}Ta	Tp2.2 presented by A10^{+}Ta
1	7.0%	ND	0.3%	81.2%
2	8.6%	ND	0.2%	0.4%
3	9.2%	ND	77.5%	0.7%
4	5.8%	ND	102.3%	0.4%
5	16.3%	ND	0.1%	93.2%
6	6.4%	ND	0.4%	84.0%
7	10.6%	ND	-0.5%	81.8%
8	12.7%	ND	77.4%	0.6%
9	25.1%	ND	0.5%	105.4%
10	28.8%	ND	0.3%	103.2%
11	30.5%	ND	-0.4%	79.4%
12	36.2%	ND	-0.3%	102.1%
13	46.7%	ND	0.5%	100.9%
14	13.9%	ND	0.1%	100.0%
15	31.6%	ND	48.7%	84.4%
16	26.5%	ND	0.0%	98.5%
17	31.1%	ND	0.5%	94.6%
18	20.5%	ND	0.3%	84.3%
19	10.6%	ND	-0.1%	68.7%
20	24.7%	ND	0.2%	99.4%
21	14.4%	ND	0.2%	83.4%
22	16.7%	ND	0.1%	86.2%
23	17.5%	ND	-0.1%	85.9%
24	28.6%	ND	0.2%	94.3%
25	16.3%	ND	1.0%	100.1%
26	18.8%	ND	0.0%	77.9%
27	10.0%	ND	2.0%	0.8%
28	11.4%	ND	-0.2%	0.5%
29	24.8%	ND	0.2%	95.5%
30	22.2%	ND	0.5%	63.9%
31	ND	ND	ND	ND
32	14.5%	ND	0.2%	76.2%

Clone Number	MHC-matched TpM	MHC-mismatched TpM	Tp2.1 presented by A10+Ta	Tp2.2 presented by A10+Ta
33	27.1%	ND	-0.1%	97.6%
34	22.9%	ND	0.0%	91.7%
35	28.7%	ND	-0.1%	94.9%
36	32.4%	ND	0.4%	86.8%
37	32.5%	ND	0.1%	91.2%
38	2.2%	ND	-0.5%	0.4%
39	30.9%	ND	-0.4%	95.9%
40	31.7%	ND	1.1%	101.1%
41	23.6%	ND	0.0%	93.5%
42	7.0%	ND	0.1%	0.0%
43	16.4%	ND	-0.2%	96.3%
44	13.3%	ND	95.9%	0.2%
45	13.4%	ND	0.5%	98.0%
46	9.2%	ND	0.1%	6.1%
47	14.3%	ND	0.1%	99.2%
48	14.8%	ND	-0.1%	94.8%
49	17.8%	ND	108.4%	1.2%
50	4.3%	ND	1.1%	0.9%
51	18.4%	ND	0.2%	88.5%
52	10.7%	ND	-0.5%	93.7%
53	19.7%	ND	-0.1%	93.6%
54	4.9%	ND	-0.1%	-0.2%
55	8.7%	ND	0.9%	98.9%
56	28.1%	ND	1.2%	101.7%
57	24.7%	ND	0.2%	105.1%
58	22.3%	ND	0.4%	94.7%
59	20.8%	ND	-0.3%	84.1%
60	20.5%	ND	0.0%	83.4%
61	15.0%	ND	99.1%	7.6%
62	30.1%	ND	0.3%	99.6%
63	49.7%	ND	0.3%	0.6%
64	14.4%	ND	1.1%	101.0%
65	9.5%	ND	101.2%	0.4%
66	14.4%	ND	0.0%	85.8%
67	3.0%	ND	0.2%	0.5%
68	28.1%	ND	0.6%	95.5%
69	39.2%	ND	0.0%	93.5%
70	20.6%	ND	0.7%	92.0%
71	21.5%	ND	-0.3%	79.3%

A) 592 clones cytotoxicity results (continued).

Clone Number	MHC-matched TpM	MHC-mismatched TpM	Tp2.1 presented by A10+Ta	Tp2.2 presented by A10+Ta
72	37.6%	ND	0.3%	94.8%
73	21.7%	ND	0.2%	121.4%
74	24.0%	ND	-0.1%	100.0%
75	15.6%	ND	0.2%	98.3%
76	30.5%	ND	0.6%	98.0%
77	33.4%	ND	0.5%	92.5%
78	10.9%	ND	0.3%	92.4%
79	30.2%	ND	0.1%	96.6%
80	15.4%	ND	0.3%	99.7%
81	18.0%	ND	90.1%	1.1%
82	25.1%	ND	0.1%	86.6%
83	19.0%	ND	-0.5%	83.3%
84	33.0%	ND	-0.1%	79.8%
85	34.0%	ND	0.1%	96.2%
86	15.3%	ND	-0.1%	97.0%
87	29.2%	ND	0.5%	98.7%
88	20.0%	ND	0.2%	101.0%
89	18.6%	ND	1.0%	96.4%
90	10.4%	ND	0.4%	86.9%
Spontaneous release 1	0.3%	ND	0.3%	0.2%
Spontaneous release 2	0.1%	ND	-0.2%	-0.3%
Spontaneous release 3	-0.1%	ND	-0.1%	0.2%

A) 592 clones cytotoxicity results (continued).

B) 1011 clones

Clone Number	MHC-matched TpM	MHC-mismatched TpM	Tp2.1 presented by A10$^+$Ta	Tp2.2 presented by A10$^+$Ta
1	25.4%	-4.7%	-5.7%	70.0%
2	74.1%	-3.6%	-5.6%	75.8%
3	62.3%	-3.0%	-4.9%	105.9%
4	27.7%	-3.3%	-5.7%	73.3%
5	69.0%	-3.4%	-5.2%	63.8%
6	65.3%	-3.8%	-5.7%	75.8%
7	15.0%	-4.1%	-5.9%	-0.1%
8	49.6%	-4.1%	-5.5%	35.0%
9	8.8%	-3.3%	-4.9%	0.2%
10	40.4%	-1.2%	-5.2%	57.7%
11	71.7%	-2.5%	-4.5%	118.1%
12	38.1%	-4.0%	-5.2%	61.4%
13	67.4%	-3.7%	-5.4%	71.0%
14	49.9%	-4.1%	-5.1%	29.9%
15	71.3%	-4.9%	-3.6%	102.3%
16	59.2%	-1.6%	-5.0%	109.6%
17	50.7%	-4.0%	-5.7%	78.8%
18	6.6%	-4.2%	-5.0%	0.8%
19	7.1%	-4.0%	-5.1%	-0.2%
20	15.9%	-3.3%	-5.5%	66.6%
21	8.3%	-4.7%	-5.0%	-0.3%
22	52.9%	-4.3%	-5.1%	90.8%
23	8.4%	-5.0%	-5.4%	0.8%
24	14.0%	-3.8%	75.1%	-0.5%
25	38.2%	-1.1%	-4.9%	83.8%
26	44.2%	-2.9%	-5.7%	84.4%
27	54.8%	-3.2%	-5.4%	87.5%
28	54.3%	-4.2%	-5.4%	76.8%
29	42.9%	5.9%	3.3%	9.0%
30	69.1%	-3.8%	-2.6%	46.1%
31	43.4%	-3.1%	-5.5%	79.0%
32	44.3%	-2.4%	-4.9%	0.0%
33	56.3%	-1.5%	-4.8%	97.3%
34	42.5%	-3.9%	-5.0%	73.3%
35	0.8%	-4.0%	-5.2%	73.9%
36	42.1%	-3.4%	-4.6%	84.4%
37	68.7%	-2.4%	-5.1%	96.2%
38	49.7%	-3.2%	-4.5%	96.6%
39	59.8%	3.2%	-3.8%	107.2%

Clone Number	MHC-matched TpM	MHC-mismatched TpM	Tp2.1 presented by A10+Ta	Tp2.2 presented by A10+Ta
40	37.4%	-3.6%	-5.2%	62.4%
41	18.5%	-1.4%	-5.2%	0.2%
42	72.6%	-3.4%	-4.6%	0.3%
43	59.1%	-2.9%	-5.2%	82.7%
44	45.2%	-4.3%	-5.6%	42.5%
45	32.2%	-0.8%	-3.7%	39.7%
46	54.1%	-3.0%	-5.6%	72.0%
47	45.3%	-4.8%	-5.8%	55.1%
48	50.6%	-4.6%	-5.4%	78.1%
49	47.9%	0.5%	-4.9%	78.6%
50	3.8%	-3.2%	-5.5%	0.1%
51	46.0%	-3.3%	-4.2%	100.5%
52	4.5%	-3.6%	-4.5%	0.5%
53	42.4%	-3.2%	-5.3%	87.6%
54	72.2%	6.7%	-3.2%	118.6%
55	8.9%	-2.8%	-4.7%	1.8%
56	7.5%	-2.1%	-4.9%	0.5%
57	10.2%	-4.7%	-4.1%	1.6%
58	32.1%	-3.7%	73.3%	-0.2%
59	78.6%	-2.2%	-3.1%	76.4%
60	66.6%	-1.5%	-5.1%	76.5%
61	60.3%	-2.4%	-0.3%	109.2%
62	71.4%	-3.1%	-4.5%	96.4%
63	19.1%	-3.7%	-3.0%	2.9%
64	18.4%	-3.2%	-5.1%	18.0%
65	2.6%	-3.9%	-5.0%	-0.7%
66	15.9%	-3.3%	5.7%	25.2%
67	36.0%	-2.1%	-4.1%	67.7%
68	9.4%	-4.2%	-4.1%	0.4%
69	60.7%	-0.6%	-4.5%	104.1%
70	ND	ND	ND	ND
71	58.3%	-3.1%	-1.1%	96.4%
72	7.3%	-3.9%	-4.6%	0.6%
73	53.5%	-3.9%	-5.0%	66.8%
74	54.3%	-4.3%	-5.0%	76.1%
75	56.1%	-3.8%	-4.8%	79.1%
76	51.4%	-4.3%	-3.9%	91.3%
77	43.1%	-3.2%	-5.3%	61.5%
78	38.6%	-3.9%	-5.6%	70.3%

B) 1011 clones cytotoxicity results (continued).

Clone Number	MHC-matched TpM	MHC-mismatched TpM	Tp2.1 presented by A10+Ta	Tp2.2 presented by A10+Ta
79	6.6%	-1.9%	-5.0%	0.9%
80	51.5%	-3.8%	-5.4%	59.7%
81	6.7%	-3.2%	-4.9%	1.3%
82	33.7%	-3.1%	-5.4%	38.4%
83	44.6%	-1.5%	-5.5%	20.1%
84	54.7%	5.2%	-1.1%	97.1%
85	52.7%	-1.1%	-3.1%	116.6%
86	60.0%	-1.9%	-4.9%	78.6%
87	8.3%	-2.8%	-4.1%	1.6%
88	64.4%	-2.3%	-3.9%	87.1%
89	26.9%	1.7%	-5.2%	49.7%
90	68.5%	1.8%	-4.3%	94.9%
Spontaneous release 1	0.5%	0.4%	3.8%	1.1%
Spontaneous release 2	-0.1%	0.9%	0.3%	0.1%
Spontaneous release 3	-0.3%	-1.3%	-4.1%	-1.2%

B) 1011 clones cytotoxicity results (continued).

C) 468 clones

Clone Number	MHC-matched TpM	MHC-mismatched TpM	$A18^{+}Ta$	Tp1 presented by $A18^{+}Ta$
1	3.5%	0.5%	ND	64.2%
2	1.7%	0.6%	ND	85.1%
3	4.4%	1.4%	ND	38.8%
4	1.1%	0.3%	ND	83.5%
5	1.1%	0.9%	ND	0.8%
6	1.5%	0.9%	ND	87.9%
7	1.6%	1.4%	ND	84.9%
8	1.5%	0.6%	ND	42.3%
9	2.9%	0.3%	ND	56.9%
10	3.6%	1.2%	ND	79.1%
11	3.0%	1.9%	ND	93.1%
12	3.7%	1.1%	ND	83.7%
13	2.2%	-1.3%	ND	-0.2%
14	7.7%	2.4%	ND	77.8%
15	1.0%	0.2%	ND	53.1%
16	1.6%	0.3%	ND	-0.1%
17	5.7%	0.0%	ND	0.0%
18	2.7%	0.8%	ND	47.2%
19	3.7%	0.1%	ND	44.9%
20	2.7%	0.3%	ND	8.0%
21	0.8%	0.0%	ND	23.6%
22	3.2%	0.6%	ND	42.0%
23	5.4%	0.3%	ND	-0.1%
24	0.2%	-0.2%	ND	-1.5%
25	2.8%	-1.3%	ND	56.8%
26	2.5%	0.6%	ND	36.1%
27	1.6%	-1.3%	ND	85.8%
28	1.6%	0.7%	ND	54.2%
29	7.5%	-0.5%	ND	125.2%
30	4.2%	0.4%	ND	-1.3%
31	0.2%	-1.1%	ND	0.3%
32	1.6%	0.0%	ND	48.2%
33	2.6%	-0.4%	ND	64.2%
34	3.8%	-0.1%	ND	72.8%
35	0.1%	-0.4%	ND	32.5%
36	2.6%	0.3%	ND	29.2%
37	4.4%	0.9%	ND	84.5%
38	1.6%	0.4%	ND	13.9%
39	0.9%	0.4%	ND	50.2%

Clone Number	MHC-matched TpM	MHC-mismatched TpM	A18+Ta	Tp1 presented by A18+Ta
40	1.4%	0.5%	ND	-1.2%
41	-0.5%	0.9%	ND	52.9%
42	2.9%	1.6%	ND	86.0%
43	1.8%	0.0%	ND	45.7%
44	1.6%	0.6%	ND	18.3%
45	3.2%	0.4%	ND	64.5%
46	2.1%	0.8%	ND	68.8%
47	1.5%	1.2%	ND	51.6%
48	1.5%	0.7%	ND	70.1%
49	2.8%	-0.4%	ND	80.3%
50	1.3%	-0.2%	ND	43.3%
51	2.2%	0.4%	ND	58.2%
52	1.1%	0.0%	ND	23.5%
53	8.0%	-0.3%	ND	84.5%
54	2.9%	-0.1%	ND	58.8%
55	1.3%	0.0%	ND	87.2%
56	0.9%	0.0%	ND	46.2%
57	1.9%	0.5%	ND	54.2%
58	0.8%	0.9%	ND	26.8%
59	1.4%	0.5%	ND	71.9%
60	1.9%	1.3%	ND	26.6%
61	2.2%	-0.7%	ND	-2.7%
62	3.2%	0.6%	ND	99.8%
63	0.5%	1.5%	ND	-0.3%
64	2.5%	0.1%	ND	74.5%
65	1.0%	0.5%	ND	59.8%
66	3.7%	0.7%	ND	100.7%
67	1.8%	0.9%	ND	105.5%
68	-0.3%	2.8%	ND	-0.5%
69	6.9%	-0.9%	ND	105.8%
70	1.7%	-0.9%	ND	-0.6%
71	3.7%	-1.0%	ND	94.2%
72	8.2%	0.7%	ND	101.1%
73	2.3%	-0.3%	ND	66.1%
74	3.3%	-0.3%	ND	57.7%
75	1.3%	0.4%	ND	74.8%
76	5.4%	0.8%	ND	86.2%
77	5.2%	0.7%	ND	106.7%
78	2.6%	1.0%	ND	1.0%

C) 468 clones cytotoxicity results (continued).

Clone Number	MHC-matched TpM	MHC-mismatched TpM	A18+Ta	Tp1 presented by A18+Ta
79	0.9%	0.5%	ND	87.5%
80	5.9%	0.4%	ND	1.9%
81	3.3%	0.3%	ND	97.4%
82	6.1%	1.2%	ND	101.1%
83	4.1%	0.5%	ND	1.3%
84	-0.3%	0.2%	ND	0.7%
85	1.2%	0.5%	ND	58.8%
86	2.1%	0.6%	ND	54.5%
87	4.3%	1.9%	ND	78.2%
88	5.3%	2.0%	ND	80.5%
89	2.1%	0.1%	ND	94.5%
90	1.1%	0.0%	ND	36.9%
Spontaneous release 1	-0.3%	0.1%	ND	0.8%
Spontaneous release 2	0.1%	-0.2%	ND	0.1%
Spontaneous release 3	0.2%	0.1%	ND	-0.9%

C) 468 clones cytotoxicity results (continued).

D) 641 clones

Clone Number	MHC-matched TpM	MHC-mismatched TpM	$A18^{+}Ta$	Tp1 presented by $A18^{+}$ Ta
1	-14.1%	-8.4%	-13.7%	-15.6%
2	19.9%	-8.7%	-10.7%	56.3%
3	14.3%	-8.7%	-14.5%	4.6%
4	8.3%	-6.1%	-11.9%	53.1%
5	4.8%	-8.0%	-14.3%	31.8%
6	-1.3%	-10.0%	-14.3%	-4.3%
7	12.7%	-7.7%	-12.3%	92.7%
8	-2.2%	-8.4%	-14.2%	15.7%
9	-0.1%	-8.2%	-14.0%	9.8%
10	-8.1%	-4.6%	-13.3%	-5.5%
11	33.3%	-7.7%	-6.1%	80.4%
12	12.9%	-5.3%	-10.9%	65.1%
13	20.7%	-9.4%	-12.6%	72.9%
14	-0.8%	-8.2%	-13.8%	88.6%
15	26.3%	-5.4%	-11.3%	110.6%
16	18.8%	-7.1%	-11.3%	96.0%
17	21.1%	-6.9%	-13.2%	91.6%
18	8.9%	-9.9%	-13.7%	112.4%
19	ND	ND	ND	ND
20	-7.4%	-8.7%	-14.4%	11.2%
21	-1.6%	-9.4%	-13.7%	0.4%
22	8.9%	-8.2%	-13.1%	53.0%
23	12.9%	-8.1%	-14.4%	46.2%
24	-9.2%	-5.4%	-14.0%	-3.2%
25	27.2%	-7.6%	-12.2%	117.2%
26	5.5%	-9.1%	-13.6%	72.7%
27	22.4%	-3.2%	-12.6%	86.6%
28	ND	ND	ND	ND
29	-8.5%	-8.1%	-14.5%	-13.2%
30	12.1%	-8.6%	-12.8%	45.2%
31	-3.9%	-9.6%	-14.3%	29.1%
32	-11.8%	-10.4%	-14.8%	-8.8%
33	1.5%	-9.6%	-14.3%	8.6%
34	2.2%	-8.1%	-14.7%	33.3%
35	-6.7%	-9.5%	-14.4%	5.4%
36	5.8%	-6.8%	-13.2%	4.6%
37	-11.3%	-10.7%	-14.4%	-18.6%
38	-19.9%	-8.7%	-13.5%	-18.0%
39	1.6%	-9.2%	-14.7%	15.7%

Clone Number	MHC-matched TpM	MHC-mismatched TpM	A18+Ta	Tp1 presented by A18+ Ta
40	14.5%	-8.0%	-14.0%	83.8%
41	26.1%	-7.4%	-13.7%	67.0%
42	-11.3%	-9.8%	-14.8%	-10.9%
43	5.3%	-6.6%	-13.7%	62.4%
44	18.1%	-8.3%	-14.1%	43.3%
45	-6.9%	-8.8%	-13.5%	9.0%
46	18.3%	-7.9%	-12.5%	70.7%
47	-6.7%	-9.8%	-14.0%	21.2%
48	2.5%	-5.2%	-11.1%	54.7%
49	ND	ND	ND	ND
50	14.7%	-10.0%	-12.4%	89.1%
51	15.9%	-8.7%	-14.1%	17.9%
52	13.8%	-7.7%	-13.3%	68.0%
53	16.4%	-9.0%	-12.5%	71.3%
54	17.0%	-9.5%	-11.2%	44.6%
55	14.7%	-7.8%	-11.1%	53.9%
56	14.3%	-10.1%	-12.9%	83.6%
57	22.3%	-8.5%	-12.3%	70.1%
58	20.7%	-7.6%	-8.3%	93.7%
59	5.8%	-10.0%	-14.8%	22.8%
60	8.9%	-8.1%	-13.3%	11.6%
61	2.6%	-4.5%	-13.8%	101.7%
62	-2.0%	-9.2%	-14.3%	4.6%
63	10.6%	-7.7%	-14.3%	9.6%
64	11.0%	-4.4%	-13.7%	83.4%
65	16.2%	-7.7%	-14.7%	73.5%
66	26.9%	-8.6%	-13.5%	90.6%
67	6.2%	-9.8%	-14.2%	42.0%
68	25.7%	-6.3%	-12.1%	78.4%
69	5.3%	-8.6%	-15.3%	59.5%
70	-3.9%	-8.2%	-15.4%	21.6%
71	-10.4%	-10.0%	-14.1%	-10.3%
72	0.7%	-9.1%	-14.6%	-2.9%
73	-2.7%	-9.4%	-14.3%	31.8%
74	-6.2%	-9.8%	-14.3%	16.2%
75	6.9%	-7.2%	-13.9%	19.4%
76	-12.9%	-8.1%	-13.3%	-18.6%
77	29.3%	-5.7%	-12.5%	82.0%
78	17.4%	-4.8%	-13.3%	85.4%

D) 641 clones cytotoxicity results (continued).

Clone Number	MHC-matched TpM	MHC-mismatched TpM	A18+Ta	Tp1 presented by A18+ Ta
79	9.2%	-4.8%	-13.8%	91.9%
80	-8.3%	-9.4%	-14.3%	6.6%
81	27.7%	-6.5%	-11.0%	31.6%
82	-2.7%	-8.6%	-12.8%	-15.6%
83	15.0%	0.2%	-12.4%	89.3%
84	9.5%	-3.8%	-12.3%	58.1%
85	-4.0%	-6.9%	-13.9%	55.2%
86	9.6%	-8.1%	-13.5%	71.6%
87	14.9%	-6.7%	-13.4%	41.7%
88	24.3%	-6.7%	-13.8%	59.4%
89	18.3%	-6.2%	-13.6%	96.9%
90	15.2%	-8.2%	-13.9%	88.8%
Spontaneous release 1	0.2%	0.8%	-0.8%	-2.0%
Spontaneous release 2	0.1%	1.4%	0.6%	-2.1%
Spontaneous release 3	-0.3%	-2.2%	0.2%	4.1%

D) 641 clones cytotoxicity results (continued).

Appendix L - Sequences of TCRβ chains expressed by clones

The sequences of the TCRβ transcripts of the A) 468, B) 592, C) 641 and D) 1011 clones. The sequences are presented in the format of Chothia *et al* (1988). Clones for which either no functional or dual functional TCRβ chains were sequenced are detailed at the foot of each table. The sequence of non-functional TCRβ chains is shown in red script and sequences for T-cells expressing dual TCRβ chain are shaded.

A) 592 clones

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
592.68 592.85	+ +	2.2 2.2	1s1	TGT GCC AGC C A S	AGC CAA GGG AGG GGG TCT GGA TTC GAG CAT S Q G R G S G F E Q	TAT TTC GGC Y F G	3s7
592.2 592.54	_	- -	1s3	TGT GCC AGC C A S	AGC CAA GAT CGA AGG GAC TCC TAC GTG CAG CTG S Q D R R D S Y V Q L	TAC TTT GGA Y F G	2s2
592.13	+	2.2	1s3	TGT GCC AGC C A S	AGC CGC GTG CAT CCG GGC AGC ATG GAG GAG CTG S R V H P G S M E E L	CAC TTC GGG H F G	3s1
592.38	-	-	1s3	TGT GCC AGC C A S	AGC CAA GGG AGA TAC AAC AAC CCT CTG S Q G R Y N N P L	TAT TTT GGA Y F G	3s3
592.63	+	-	1s3	TGT GCC AGC C A S	AGC CGA GGG GGG CGA GGC AAC CCT CTG S R G G R G N P L	TTC TCC GGA Y F G	3s3
592.1 592.41 592.75	- + +	2.2 2.2 2.2	1s6	TGT GCC AGC C A S	AGC CCG AAT TCC TAT GAG CAG S P N S Y E Q	TAT TTC GGC Y F G	3s7
592.42	-	-	1s7	TGT GCC AGC C A S	AGC CAA GAT ATA GGG GGT TCA ACT CAG ATC CAG S Q D I G G S T Q I Q	TAC TTC GGG Y F G	3s5
592.67	-	-	1s13	TGT GCC AGC C A S	AGT GAA TCT CGT GGG GGC GGA TAT GAG CAG S E S R G G G Y E Q	TAT TTC GGC Y F G	3s7

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
592.7 592.15 592.18 592.51 592.70 592.78 592.87	+ + + + + + +	2.2 2.1/2.2 2.2 2.2 2.2 2.2 2.2	2s4	TGC AGT GCT C S A	CAA TGG GGG GGT TCC TAT GAG GAG CAG Q W G G S Y E E Q	CAC TTC GGG H F G	2s1
592.27	+	-	6s2	TGT GCC AGC C A S	CTC GCC AGC AGG AAC ACT GAG GTT L A S R N T E V	TTC TTT GGA F F G	1s1
592.34 592.90	+ +	2.2 2.2	6s2	TGT GCC AGC C A S	AGC CCC ATC TTG GCT TGG GAG ACG CTG S P I L A W E T L	TAC TTC GGG Y F G	2s4
592.21	+	2.2	6s3	TGT GCC AGC C A S	AGC CCA ACT CGG ACT TCG ACT AAC ACG CAG S P T R T S T N T Q	TAC TTC GGC Y F G	3s4
592.46	-	-	8s1	TGT GCC AGC C A S	AGC AAG AGC ACT GAC ACT GAG GTT S K S T D T E V	TTC TTT GGA F F G	1s1
592.14 592.86	+ +	2.2 2.2	8s2	TGT GCC AGC C A S	AGC AAA TCG TTG TTT GAG GAT CCA S K S L F E I Q	TAC TTC GG Y F G	3s5
592.52	+	2.2	8s2	TGT GCC AGC C A S	AGC TAT GAG ACT TCG GGA TAT GGG GAG CTG S Y E T S G Y G E L	CAC TTC GGG H F G	3s1
592.46			10s6	TGT GCC AGC C A S	AGC CAT CAG GAT AGG GAT GAC TAT S H Q D R D D Y	CAC TTC GGC H F G	1s2
592.12 592.43 592.48 592.55	+ + + -	2.2 2.2 2.2 2.2	12s1	TGC GCC AGC C A S	CAC ATT CGG GGG GGG CTC GAC ACC CAG CCC CTG H I R G G L D T Q P L	TAC TTT GGA Y F G	3s2
592.16	+	2.2	13s2	TGT GAA AGC C A S	AGC CAC GCG GGA TAT GAG CAG S H A G Y E Q	TAT TTC GGC Y F G	3s7

A) Sequences of 592 clones (continued).

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
592.58 592.73	+ +	2.2 2.2	13s4	TGC GCC AGC C A S	AGT TCA TCC CTC CTA GAC ACG CAG S S S L L D T Q	TAC TTC GGC Y F G	3s4
592.21	+	2.2	13s5	TGT ACC AGC C T S	AGT TAC TCT GGC AGC ATA ACG CTA CAG GAC TAT S Y S G S I T L Q D Y	CAC TTC GGC H F G	1s2
592.1 592.9 592.10 592.17 592.23 592.24 592.25 592.29 592.35 592.45 592.53 592.64 592.72 592.76 592.79 592.84 592.89			13s6	AGC GCC TGC S A C	GGT TAT GAT CGG GAC AGG TAT GGG GAG CTG G Y D R D R Y G E L	CAC TTC GGG H F G	3s1
592.63			13s6	AGC GCC TGC S A C	GGT TTC ACC CAC AGC GGC CTA GAC GCG ACG ATA AAT TCG CCC CTC G F T H S G L D A T I N S P L	CAC TTT GGA H F G	1s5
592.5 592.6 592.36 592.39 592.59	+ - + + +	2.2 2.2 2.2 2.2 2.2	13s7	TGT GCC AGC C A S	GGC CAT CTT GTC GGG GGG ACG AGG GAG CTG G H L V G G T R E L	CAC TTC GGG H F G	3s1
592.56	+	2.2	13s8	TGT GCC AGC C A S	AGT CAG GGA CAG CAG GGC TAT GTG AAC AAC CCT CTG S Q G Q Q G Y V N N P L	TAT TTT GGA Y F G	3s3

A) Sequences of 592 clones (continued).

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
592.20 592.83	+ +	2.2 2.2	13s9	TGC GCC AGC C A S	GGC GAG CAG G E Q	TAT TTC GGC Y F G	3s7
592.3 592.4 592.8 592.15 592.44 592.49 592.61 592.65 592.81	- - + + + + + - +	2.1 2.1 2.1 2.1/2.2 2.1 2.1 2.1 2.1 2.1	14s1	TGT GCC AGC C A S	AGT GTC TCC TTC GGG GGG GCC CCT TAT GGG GAG CTG S V S F G G A P Y G E L	CAC TTC GGG H F G	3s1
592.26 592.41 592.66 592.72 592.74 592.80	+ + + + + +	2.2 2.2 2.2 2.2 2.2 2.2	14s1	TGT GCC AGC C A S	AGT GAA TCG CAA GGG GCG AAT TAT GAC TAT S E S Q G A N Y D Y	CAC TTC GGC H F G	1s2
592.11 592.22 592.30 592.32 592.62 592.71	+ + + + + +	2.2 2.2 2.2 2.2 2.2 2.2	17s1	TGT GCT GCT C A A	AGT TCT GTC TTC GGG GGG GAC CTT AAA TCC TAT GAG CAG S S V F G G D L K S Y E Q	TAT TTC GGC Y F G	3s7
592.63	+	-	17s1	TGT GCT GCT C A A	AGT AGG CCT TTA GGA CAC CTC AAC AAC CCT CTG S R P L G H L N N P L	TAT TTT GGA Y F G	3s3
592.50	-	-	17s2	TGC TCT GGG C S G	GCC TCC GGG CTA GAC TAT A S G L D Y	CAC TTC GGC H F G	1s2
592.84	+	2.2	17s3	TGT GCT GAT C A A	TGG GAC AGC GGG AGC GAC TAT W D S G S D Y	CAC TTC GGC H F G	1s2
592.58* 592.73*			17s3	TGT GCT GCC C A A	CAC CCG AGA CTG GGG GGT GTA A AGA CAC CTG TAT GAG GAG CAG	CAC TTC GGG H F G	2s1

A) Sequences of 592 clones (continued).

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
592.28	+	-	17s3	TGT GCT GCT C A A	AGT AGA GAT ATG ACA GGC ACT GAG GTT S R D M T G T E V	TTC TTT GGA F F G	1s1
592.1	-	2.2	28s1	TGT GCC AGC C A S	GCT GAA TAT GGG GGG GAG AAC ACC CAG CCC CTG A E Y G G E N T Q P L	TAC TTT GGA Y F G	3s2
592.9	+	2.2					
592.10	+	2.2					
592.17	+	2.2					
592.23	+	2.2					
592.24	+	2.2					
592.25	+	2.2					
592.29	+	2.2					
592.35	+	2.2					
592.45	+	2.2					
592.53	+	2.2					
592.64	+	2.2					
592.72	+	2.2					
592.76	+	2.2					
592.79	+	2.2					
592.84	+	2.2					
592.89	+	2.2					
592.33	+	2.2	28s1	TGT GCC AGC C A S	GCT GAA TAT GGG GGG GAG AAC ACC CAG CCC CTG A E Y G G E N T Q P L	TAC TTT GGA Y F G	3s2
592.37	+	2.2					
592.13	+	2.2	28s1	TGT GCC AGC C A S	GGG GGC CGG GAC AGC ATA TAT GAC TAT G G R D S I Y D Y	CAC TTC GGC H F G	1s2
592.19	+	2.2					
592.40	+	2.2					
592.47	+	2.2					
592.57	+	2.2					
592.69	+	2.2					
592.77	+	2.2					

A) Sequences of 592 clones (continued).

Clones with dual functional TCRβ chains sequenced (8) - 592.1, 13, 15, 21, 41, 63, 72 and 84

Clones with no functional TCRβ chains sequenced (3) - 592.60, 82 and 88

B) 1011 clones

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
1011.55	-	-	1s3	TGT GCC AGC C A S	AGT CCA GCA GGG AGG GAC ACG CAG S P A G R D T Q	TAT TTC GGC Y F G	3s4
1011.1 1011.39 1011.73	+ + +	2.2 2.2 2.2	1s7	TGT GCC AGC C A S	AGC CAA GTG GGG GGG ATC TAT GGG GAG CTG S Q V G G I Y G E L	CAC TTC GGG H F G	3s1
1011.30 1011.88	+ +	2.2 2.2	1s7	TGT GCC AGC C A S	AGC CCT AGA CTT CGG GAT AGC AAC AAC CCT CTG S P R L R D S N N P L	TAT TTT GGA Y F G	3s3
1011.6	+	2.2	1s8	TGT GCC AGC C A S	AGT CCG GGA CAC GAG GAG CAG S P G H E E Q	CAC TTC GGG H F G	2s1
1011.56	-	-	1s11	TGT GCC AGC C A S	AGC TCA GGC GCA GCT AAT AGC AAC AAC CCT CTG S S G A A N S N N P L	TAT TTT GGA Y F G	3s3
1011.57	+	-	2s3	TGC AGT GCT C S A	TCG ATG GGG TTC AGG GGC CAA AGC ACT CAG S M G F R G Q S T Q	TAC TTC GGC Y F G	2s3
1011.68	-	-	2s3	TGC AGT GCT C S A	AGT TGG GGG GTG GGC GGC CAA AGC ACT CAG S W G V G G Q S T Q	TAC TTC GGC Y F G	2s3
1011.37	+	2.2	2s4	TGC AGT GCT C S A	AGA TTT GGC CCG GGG GGG TTA TCC TAT GAG CAG R F G P G G L S Y E Q	TAT TTC GGC Y F G	3s7
1011.24	+	2.1	4s3	TGC AGT GCT C S A	CCT GGA CAG CAG GGC TAT GAG CAG P G Q Q G Y E Q	TAT TTC GGC Y F G	3s7
1011.79	-	-	4s3	TGC AGT GCA C S A	GTG ATA ACG GTA GAC ACG CAG V I T V D T Q	TAC TTC GGC Y F G	3s4
1011.22 1011.47 1011.49 1011.75 1011.88	+ + + + +	2.2 2.2 2.2 2.2 2.2	4s6	TGC AGT GCT C S A	GGA GTT GGC ACC CGG CTC TCC TAT GAG CAG G V G T R L S Y E Q	TAT TTC GGC Y F G	3s7

B) Sequences of 1011 clones (continued).

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
1011.18	-	-	5s1	TGT GCC AGC C A S	AGC CCG ACT TCG GGC CCG AGC AAC AAC CCT CTG S P T S G P S N N P L	TAT TTT GGA Y F G	3s3
1011.52	-	-	6s2	TGT GCC AGC C A S	AGT CTA CAC GGT AAC ACT GAG GTT S L H G N T E V	TTC TTT GGA F F G	1s1
1011.7	+	-	6s3	TGT GCC AGC C A S	AGC CCA ACG GGG GGG GGA TCA GCC CAC AAC AAC CCT CTG S P T G G G S A H N N P L	TAT TTT GGA Y F G	3s3
1011.50	-	-	10s3	TGC GCC AGC C A S	AGC CAT AGG ACA GCT AGT TAT GAC TAT S H R T A S Y D Y	CAC TTC GGC H F G	1s2
1011.80	+	2.2	12s1	TGC GCC AGC C A S	AGC CTC GTG GGG AGA GAG ACG CTG S L V G R E T L	TAC TTC GGG Y F G	2s4
1011.12 1011.27	+ +	2.2 2.2	12s2	TGC GCC AGC C A S	AGT TAC TCG CCC GGG GGG GGG TCG CCT CTG S Y S P G G G S P L	TAT TTT GGA Y F G	3s3
1011.63	+	-	13s1	TGT GCC AGC C A S	AGT TAC CAG GGA CCG AAC CCT CTG S Y Q G P N P L	TAT TTT GGA Y F G	3s3
1011.71	+	2.2	13s1	TGT GCC AGC C A S	AGT AAC AGC AGG GCC ACC GCT TAT GAC TAT S N S R A T A Y D Y	CAC TTC GGC H F G	1s2
1011.23	-	-	13s2	TGT GAA AGC C A S	AGT TAT CAG CCG ATC AAT TCG CCC CTC S Y Q P I N S P L	CAC TTT GGA H F G	1s5
1011.42	+	-	13s4	TGC GCC AGA C A R	GGA CTG GGG TAT CGG GAG ACG CTG G L G Y R E T L	TAC TTC GGG Y F G	2s4
1011.58	+	2.1	13s5	TGT ACC AGC C T S	AGT CGC GGG GGG CGC ATT GAT GGG GAG CTG S R G G R I D G E L	CAC TTC GGG H F G	3s1

B) Sequences of 1011 clones (continued).

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
1011.2	+	2.2	14s1	TGT GCC AGC C A S	AGT GTT GGT AAC AGC AAC TAT GAG CAG S V G N S N Y E Q	TAT TTC GGC Y F G	3s7
1011.4	+	2.2					
1011.6	+	2.2					
1011.10	+	2.2					
1011.13	+	2.2					
1011.14	+	2.2					
1011.20	+	2.2					
1011.25	+	2.2					
1011.26	+	2.2					
1011.28	+	2.2					
1011.30	+	2.2					
1011.31	+	2.2					
1011.34	+	2.2					
1011.35	-	2.2					
1011.36	+	2.2					
1011.44	+	2.2					
1011.45	+	2.2					
1011.46	+	2.2					
1011.51	+	2.2					
1011.53	+	2.2					
1011.59	+	2.2					
1011.60	+	2.2					
1011.66	+	2.2					
1011.67	+	2.2					
1011.74	+	2.2					
1011.76	+	2.2					
1011.77	+	2.2					
1011.78	+	2.2					
1011.82	+	2.2					
1011.83	+	2.2					
1011.89	+	2.2					
1011.9	-	-	15s2	TGT GCC AGC C A S	AAT AGC GGA CAG CAG GGT GGC ACC CAG CCC CTG N S G Q Q G G T Q P L	TAC TTT GGA Y F G	3s2
1011.19	-	-	15s2	TGT GCC AGC C A S	AAC CCC GGA CAG CAG GGA GGC ACC CAG CCC CTG N P G Q Q G G T Q P L	TAC TTT GGA Y F G	3s2

B) Sequences of 1011 clones (continued).

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
1011.21	-	-	15s2	TGT GCC AGC C A S	AAC CCG GGA CAG CAG GGA GGC GCC GTG CAG CTG N P G Q Q G G A V Q L	TAC TTT GGA Y F G	2s2
1011.72 1011.81	- -	- -	15s2	TGT GCC AGC C A S	AAC CGA GGA CAA CAG GGG GGA TCC TAT GAG CAG N R G Q Q G G S Y E Q	TAT TTC GGC Y F G	3s7
1011.59 1011.65	+ -	2.2 -	16s1	TGT GCC AGC C A S	AGC CGA GGT GAG GCA AAC TCA GAG ACG CTG S R G E A N S E T L	TAC TTC GGG Y F G	2s4
1011.32	+	-	20s1	TGT GCC TTA C A L	AGG GCT GGG GGG TCT TCA GAG ACG CTG R A G G S S E T L	TAC TTC GGG Y F G	2s4
1011.40	+	2.2	28s1	TGT GCC AGC C A S	GCT GAA TAC GGG GGC GAG AAC ACC CAG CCC CTG A E Y G G E N T Q P L	TAC TTT GGA Y F G	3s2
1011.5 1011.8 1011.11 1011.15 1011.16 1011.17 1011.33 1011.38 1011.43 1011.54 1011.61 1011.62 1011.64 1011.69 1011.84 1011.85 1011.86 1011.90	+ + + + + + + + + + + + + + + + + +	2.2 2.2 2.2 2.2 2.2 2.2 2.2 2.2 2.2 2.2 2.2 2.2 2.2 2.2 2.2 2.2 2.2 2.2	X	TGC ACC TGC C T C	AGT AAA GCG GCA GCT GAG GAC GGT TAT GAG CAG S K A A A E D G Y E Q	TAT TTC GGC Y F G	3s7
1011.3	+	2.2	X	TGC ACC TGC C T C	AGT CGC GGT GAT AAT TCC TAT GAG CAG S R G D N S Y E Q	TAT TTC GGC Y F G	3s7

Clones with dual functional TCRβ chains sequenced (4) - 1011.6, 30, 59 and 88

Clones with no functional TCRβ chains sequenced (4) - 1011.29, 41, 48 and 87

C) 468 clones

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
468.35	-	1	1s7	TGT GCC AGC	AGC GAT GAC TTT TAT AGC ACA GAC ACG CAG	TAC TTC GGC	3s4
468.52	-	1		C A S	S D D F Y S T D T Q	Y F G	
468.12	-	1	1s9	TGT GCC AGC	AGC CAA GAC TAC CCA ACC AAC GAC CCT CTG	TAT TTT GGA	3s3
468.15	-	1		C A S	S Q D Y P T N D P L	Y F G	
468.27	-	1					
468.29	-	1					
468.33	-	1					
468.39	-	1					
468.57	-	1					
468.74	-	1					
468.79	-	1					
468.40	-	-	2s7	TGT GGT GCT	AGA GCT CCT CGG GGT CCA TCC CGG GAG	TAT TTT GGA	3s3
468.68	-	-		C G A	R A P R G P S R E	Y F G	
468.44	-	1	2s7	TGT GGT GTT	CCA TCT GGG ACA GCA AGT GAC TAT	CAC TTT GGC	1s2
468.55	-	1		C G V	P S G T A S D Y	H F G	
468.75	-	1	2s9	TGT GGT GCC	CAT GGA CCC GAT TAC TCC ACA GAC ACG CAG	TAC TTC GGC	3s4
				C G A	H G P D Y S T D T Q	Y F G	
468.1	-	1	3s1	TGC GCC AGC	AGC TAT TTC GGG GTC ACC CAG CCC CTG	TAC TTT GGA	3s2
468.4	-	1		C A S	S Y F G V T Q P L	Y F G	
468.9	-	1					
468.19	-	1					
468.28	-	1					
468.34	-	1					
468.36	-	1					
468.41	-	1					
468.45	-	1					
468.46	-	1					
468.48	-	1					
468.50	-	1					
468.51	-	1					
468.53	-	1					

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
468.56 468.77 468.79 468.82 468.85 468.86 468.90	- - - - - - -	1 1 1 1 1 1 1	3s1	TGC GCC AGC C A S	AGC TAT TTC GGG GTC ACC CAG CCC CTG S Y F G V T Q P L	TAC TTT GGA Y F G	3s2
468.6	-	1	6s2	TGT GCC AGC C A S	AGC CGA TCT TCG GGG CAT GTT AGC AAC GAC CCT CTG S R S S G H V S N D P L	TAT TTT GGA Y F G	3s3
468.3 468.7 468.11 468.14 468.18 468.21 468.22 468.26 468.38 468.42 468.47 468.49 468.59 468.66 468.67 468.71 468.76 468.87 468.88 468.89	- 	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	10s4	TGT GCC AGC C A S	AGG CGA GCG GTT ATG AAC GAC CCT CTG R R A V M N D P L	TAT TTT GGA Y F G	3s3

C) Sequences of 468 clones (continued).

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
468.16 468.17 468.30 468.32 468.78 468.80 468.83 468.84	- - - - - - - -	- - - 1 - - - -	13s2	TGT GCC AGC C A S	AGT TAT GAG GAG CAC GGA GAG CAG S Y E E H G E Q	TAT TTC GGC Y F G	3s7
468.37 468.54	- -	1 1	13s1	TGT GCC AGC C A S	AGA GGG GGG GCA GGG ACG AAT GGG GAG CTG R G G A G R N G E L	CAC TTC GGG H F G	3s1
468.25	-	1	13s1	TGT GCC AGC C A S	AGG ATC ACA GCA GGG GAG CAG R I T A G E Q	CAC TTC GGG H F G	2s1
468.5	-	-	14s1	TGT GCC AGC C A S	TCT AAG GAC TCC GTC CCA GAG AGA CTC S K D S V P E R L	TAT TTC GGC Y F G	3s3
468.41 468.43	- -	1 1	14s1	TGT GCC AGC C A S	AGT GCC ACG ACA GCA GGG CTT GCA GGC GCC GCC CTG S A T T A G L A G A A L	ACC TTC GGG T F G	3s6
468.21 468.89	- -	1 1	15s1	TGT GCC AGC C A S	AAT GGG GCC TCA AAC GAC CCT CTG N G A S G D P L	TAT TTT GGA Y F G	3s3
468.38	-	1	17s3	TGT GCT GCT C A A	GAA GCC TTC GGG GGG CCG CTC GAC GAC CCT CTG E A F G G P L D D P L	TAT TTT GGA Y F G	3s3
468.2	-	1	20s1	TGT GCC TGG C A W	AGT CGC GGG AAT GAG CAG S R G N E Q	TAT TTC GGC Y F G	3s7
468.20	-	-	20s1	TGT GCC TGG C A W	GAG TCC GAG GAG ACG CTG E S E E T L	TAC TTC GGG Y F G	2s4
468.24	-	-	20s1	TGT GCC TGG C A W	ACT CAC CAG ACG GAG AGT TAT GAC TAT T H Q T E S Y D Y	CAC TTT GGC H F G	1s2

C) Sequences of 468 clones (continued).

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
468.37	-	1	20s1	TGT GCC TGG C A W	ATC AGC GAG TAC GAC TAT I S E Y D Y	CAC TTT GGC H F G	1s2
468.8 468.32	- -	1 1	24s1	TGT GCC AGC C A S	AGA GCG CGG ACT CCG AGC AAC GAC CCT CTG R A R T P S N D P L	TAT TTT GGA Y F G	3s3

C) Sequences of 468 clones (continued).

Clones with dual functional TCRβ chains sequenced (7) - 468. 21, 32, 37, 38, 41, 79 and 89

Clones with no functional TCRβ chains sequenced (16) - 468.10, 13, 23, 31, 58, 60, 61, 62, 63, 64, 65, 69, 70, 72, 73 and 81

D) 641 clones

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
641.10	-	-	1s7	TGT GCC AGC	AGC CAC GAG TGG TAC AGC ACA GAC ACG CAG	TAC TTC GGC	3s4
641.14	-	1		C A S	S H E W Y S T D T Q	Y F G	
641.15	+	1					
641.16	+	1					
641.18	-	1					
641.21	-	-					
641.31	-	1					
641.33	-	-					
641.35	-	-					
641.43	-	1					
641.46	+	1					
641.52	+	1					
641.61	-	1					
641.63	+	-					
641.64	+	1					
641.66	+	1					
641.69	-	1					
641.73	-	1					
641.74	-	1					
641.79	-	1					
641.81	+	1					
641.83	+	1					
641.85	-	1					
641.62	-	-	1s9	TGT GCC AGC	AGC CCA GAC ATC CCG ACG ATC GGT GAG CGG	TAT TTT GGC	2s5
641.80	-	-		C A S	S P D I P T I G E R	Y F G	
641.47	-	1	1s9	TGT GCC AGC	AGC CAA GAC TAC CCA ACC AAC GAC CCT CTG	TAT TTT GGA	3s3
				C A S	S Q D Y P T N D P L	Y F G	

D) Sequences of 641 clones (continued).

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
641.1	-	-	3s1	TGC GCC AGC	AGG GAA AAA GGA CTT CGG GGG TCT AGC AAC GAC CCT CTG	TAT TTT GGA	3s3
641.2	+	1		C A S	R E K G L R G S S N D P L	Y F G	
641.3	+	-					
641.5	-	1					
641.6	-	-					
641.7	+	1					
641.9	-	-					
641.11	+	1					
641.13	+	1					
641.17	+	1					
641.22	-	1					
641.23	+	1					
641.26	-	1					
641.27	+	1					
641.29	-	-					
641.30	+	1					
641.32	-	-					
641.35	-	-					
641.40	+	1					
641.42	-	-					
641.44	+	1					
641.48	-	1					
641.50	+	1					
641.51	+	1					
641.53	+	1					
641.54	+	1					
641.55	+	1					
641.56	+	1					
641.57	+	1					
641.58	+	1					
641.59	-	1					
641.60	-	1					
641.65	+	1					
641.67	-	1					
641.68	+	1					
641.70	-	1					
641.71	-	-					

D) Sequences of 641 clones (continued).

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
641.72 641.75 641.77 641.78 641.84 641.87 641.89	- - + + - + +	- 1 1 1 1 1 1	3s1	TGC GCC AGC C A S	AGG GAA AAA GGA CTT CGG GGG TCT AGC AAC GAC CCT CTG R E K G L R G S S N D P L	TAT TTT GGA Y F G	3s3
641.76	-	-	3s1	TGC GCC AGC C A S	AGT GCG AAG GCC CAC GGC AGG GTC CAT GCA GGC GCC GCC CTG S A K A H G R V H A G A A L	ACC TTC GGG T F G	3s6
641.12	+	1	3s1	TGC GCC AGC C A S	AAT AAA TGG GGG GGC GGC AAC AAC CCT CTG N K W G G G N N P L	TAT TTT GGA Y F G	3s3
641.8	-	1	4s5	TGC AGT GCT C S A	AGG CCA GAC TTC GGA GGA TTC AAC GAC CCT CTG R P D F G G F N D P L	TAT TTT GGC Y F G	3s3
641.39	-	1	4s5	TGC TAC GGT C S A	AGG TCC GGA GCT TGG GAC TAT R S G A W D Y	CAC TTT GGC H F G	1s2
641.82	-	-	7s1	TGC GCC AGC C A S	AGT AGA GAG CCT CAA GAC CGA GAG ATA S R E P Q D R E I	CTA TCG GGG L F G	2s1
641.20	-	1	10s7	TGT GCC AGC C A S	AGC CAG AGA TAC GAT GAG CAG S Q R Y D E Q	TAT TTC GGC Y F G	3s7
641.37	-	-	13s10	TGC GCC AGC C A S	AGG GGC GAA TCC CAC AAC CCT CTG R G E S H N P L	TAT TTT GGA Y F G	3s3
641.25 641.86 641.90	+ - +	1 1 1	14s1	TGT GCC AGC C A S	AGT TCT AGC TGG GAG CAG S S S W E Q	TAT TTC GGC Y F G	3s7
641.36 641.38	- -	- -	14s1	TGT GCC AGC C A S	AGC CGG GGC TGG GAG CAG S R G W E Q	TAT TTC GGC Y F G	3s7

D) Sequences of 641 clones (continued).

T-cell clone	Cytotoxicity		Sequence				
	TpM	Epitope	Vβ	FR	CDR3	FR	Jβ
641.34	-	1	14s1	TGT GCC AGC C A S	TCC GGG GGA CAG CTT AGT GAG CGG S G G Q L S E R	TAT TTT GGC Y F G	2s5
641.41	+	1	14s1	TGT GCC AGC C A S	AGT GAA TCG TGG GAG CTG S E S W E L	CAC TTC GGG H F G	2s1
641.88	+	1	14s1	TGT GCC AGC C A S	AGT GAT CAC CGG GCA GGG CCG TGG GAG CAG S D H R A G P W E Q	TAT TTC GGC Y F G	3s7
641.4	-	1	16s1	TGT GCC AGC C A S	AGC CTT GGG GCC CTT GAG GTT S L G A L E V	TTC TTT GGA F F G	1s1

D) Sequences of 641 clones (continued).

Clones with dual functional TCRβ chains sequenced (1) - 641.35

Clones with no functional TCRβ chains sequenced (2) - 641.24 and 45

9 References

Aandahl, E. M., Sandberg, J. K., Beckerman, K. P., Tasken, K., Moretto, W. J. and Nixon, D. F. (2003). CD7 Is a differentiation marker that identifies multiple CD8+ T-cell effector subsets. *J Immunol* **170**, 2349-2355.

Akamatsu, Y., Tsurushita, N., Nagawa, F., Matsuoka, M., Okazaki, K., Imai, M. and Sakano, H. (1994). Essential residues in V(D)J recombination signals. *J Immunol* **153**, 4520-9.

Akolkar, P. N., Gulwani-Akolkar, B., Pergolizzi, R., Bigler, R. D. and Silver, J. (1993). Influence of HLA genes on T-cell receptor V segment frequencies and expression levels in peripheral blood lymphocytes. *J Immunol* **150**, 2761-73.

Allen, T. M., O'Connor, D. H., Jing, P., Dzuris, J. L., Mothe, B. R., Vogel, T. U., Dunphy, E., Liebl, M. E., Emerson, C., Wilson, N. *et al.* (2000). Tat-specific cytotoxic T lymphocytes select for SIV escape variants during resolution of primary viraemia. *Nature* **407**, 386-390.

Allsopp, B. A. and Allsopp, M. T. (1988). *Theileria parva*: genomic DNA studies reveal intra-specific sequence diversity. *Mol Biochem Parasitol* **28**, 77-83.

Allsopp, B., Carrington, M., Baylis, H., Sohal, S., Dolan, T. and Iams, K. (1989). Improved characterization of *Theileria parva* isolates using the polymerase chain reaction and oligonucleotide probes. *Mol Biochem Parasitol* **35**, 137-47.

Altfeld, M. and Allen, T. M. (2006). Hitting HIV where it hurts: an alternative approach to HIV vaccine design. *Trends in Immunology* **27**, 504-510.

Anderson, S. J., Chou, H. S. and Loh, D. Y. (1988). A conserved sequence in the T-cell receptor beta-chain promoter region. *PNAS* **85**, 3551-3554.

Annels, N. E., Callan, M. F. C., Tan, L. and Rickinson, A. B. (2000). Changing patterns of dominant TCR usage with maturation of an EBV-specific cytotoxic T-cell response. *J Immunol* **165**, 4831-4841.

Anton, L. C., Yewdell, J. W. and Bennink, J. R. (1997). MHC class I-associated peptides produced from endogenous gene products with vastly different efficiencies. *J Immunol* **158**, 2535-42.

Antonacci, R., Massari, S., De Iaco, R. and Ciccarese, S. (2001). Assignment of the TRB@ locus encoding the T-cell receptor beta chain to sheep, cattle, goat and river buffalo chromosomes by in situ hybridization. *Cytogenet Cell Genet* **94**, 82-3.

Arden, B., Clark, S. P., Kabelitz, D. and Mak, T. W. (1995a). Human T-cell receptor variable gene segment families. *Immunogenetics* **42**, 455-500.

Arden, B., Clark, S. P., Kabelitz, D. and Mak, T. W. (1995b). Mouse T-cell receptor variable gene segment families. *Immunogenetics* **42**, 501-530.

Argaet, V. P., Schmidt, C. W., Burrows, S. R., Silins, S. L., Kurilla, M. G., Doolan, D. L., Suhrbier, A., Moss, D. J., Kieff, E. and Suclley, T. B. (1994). Dominant selection of an invariant T-cell antigen receptor in response to persistent infection by Epstein-Barr virus. *J. Exp. Med.* **180**, 2335-2340.

Arstila, T. P., Casrouge, A., Baron, V., eacute, ronique, Even, J., Kanellopoulos, J. and Kourilsky, P. (1999). A direct estimate of the human T-cell receptor diversity. *Science* **286**, 958-961.

Attuil, V., Bucher, P., Rossi, M., Mutin, M. and Maryanski, J. L. (2000). Comparative T-cell receptor repertoire selection by antigen after adoptive transfer: a glimpse at an antigen-specific preimmune repertoire. *Proc Natl Acad Sci U S A* **97**, 8473-8.

Baldwin, C. L., Goddeeris, B. M. and Morrison, W. I. (1987). Bovine helper T-cell clones specific for lymphocytes infected with *Theileria parva* (Muguga). *Parasite Immunol* **9**, 499-513.

Baldwin, C. L., Black, S. J., Brown, W. C., Conrad, P. A., Goddeeris, B. M., Kinuthia, S. W., Lalor, P. A., MacHugh, N. D., Morrison, W. I., Morzaria, S. P. *et al.* (1988). Bovine T-cells, B-cells, and null cells are transformed by the protozoan parasite *Theileria parva*. *Infect Immun* **56**, 462-7.

Baldwin, C. L., Iams, K. P., Brown, W. C. and Grab, D. J. (1992). *Theileria parva:* CD4+ helper and cytotoxic T-cell clones react with a schizont-derived antigen associated with the surface of *Theileria parva*-infected lymphocytes. *Exp Parasitol* **75**, 19-30.

Ballingall, K. T., Mwangi, D. M., MacHugh, N. D., Taracha, E. L. N., Totte, P. and McKeever, D. J. (2000). A highly sensitive, non-radioactive assay for T-cell activation in cattle: applications in screening for antigens recognised by CD4+ and CD8+ T-cells. *Journal of Immunological Methods* **239**, 85-93.

Barker, P. E., Ruddle, F. H., Royer, H. D., Acuto, O. and Reinherz, E. L. (1984). Chromosomal location of human T-cell receptor gene Ti beta. *Science* **226**, 348-9.

Baron, C., Sachs, D. H. and LeGuern, C. (2001). A particular TCR beta variable region used by T cells infiltrating kidney transplants. *J Immunol* **166**, 2589-96.

Barouch, D. H., Kunstman, J., Kuroda, M. J., Schmitz, J. E., Santra, S., Peyerl, F. W., Krivulka, G. R., Beaudry, K., Lifton, M. A., Gorgone, D. A. *et al.* (2002). Eventual AIDS vaccine failure in a rhesus monkey by viral escape from cytotoxic T lymphocytes. *Nature* **415**, 335-9.

Basler, M., Youhnovski, N., Van Den Broek, M., Przybylski, M. and Groettrup, M. (2004). Immunoproteasomes down-regulate presentation of a subdominant T-cell epitope from lymphocytic choriomeningitis virus. *J Immunol* **173**, 3925-34.

Bassing, C. H., Alt, F. W., Hughes, M. M., D'Auteuil, M., Wehrly, T. D., Woodman, B. B., Gartner, F., White, J. M., Davidson, L. and Sleckman, B. P. (2000). Recombination signal sequences restrict chromosomal V(D)J recombination beyond the 12/23 rule. *Nature* **405**, 583-6.

Bassing, C. H., Swat, W. and Alt, F. W. (2002). The mechanism and regulation of chromosomal V(D)J recombination. *Cell* **109**, S45-S55.

Beekman, N. J., van Veelen, P. A., van Hall, T., Neisig, A., Sijts, A., Camps, M., Kloetzel, P. M., Neefjes, J. J., Melief, C. J. and Ossendorp, F. (2000). Abrogation of CTL epitope processing by single amino acid substitution flanking the C-terminal proteasome cleavage site. *J Immunol* **164**, 1898-905.

Behlke, M. A., Chou, H. S., Huppi, K. and Loh, D. Y. (1986). Murine T-cell receptor mutants with deletions of beta-chain variable region genes. *Proc Natl Acad Sci U S A* **83**, 767-71.

Belz, G. T., Stevenson, P. G. and Doherty, P. C. (2000). Contemporary analysis of MHC-related immunodominance hierarchies in the CD8+ T-cell response to influenza A viruses. *J Immunol* **165**, 2404-9.

Bennett, A. R., Gorak-Stolinska, P., Ben-Smith, A., Floyd, S., de Lara, C. M., Weir, R. E., Lalor, M. K., Makamo, K., Msiska, G. K., Crampin, A. C. and Beverley, P. (2006). The PPD-specific T-cell clonal response in UK and Malawian subjects following BCG vaccination: A new repertoire evolves over 12 months. *Vaccine* **24**, 2617-2626.

Benoist, C. and Mathis, D. (1989). Positive selection of the T-cell repertoire: where and when does it occur? *Cell* **58**, 1027-33.

Bensaid, A. and Hadam, M. (1991). Bovine CD4 (BoCD4). *Veterinary Immunology and Immunopathology* **27**, 51-54.

Bernardin, F., Doukhan, L., Longone-Miller, A., Champagne, P., Sekaly, R. and Delwart, E. (2003). Estimate of the total number of CD8+ clonal expansions in healthy adults using a new DNA heteroduplex-tracking assay for CDR3 repertoire analysis. *Journal of Immunological Methods* **274**, 159-175.

Bernardin, F., Magierowska, M., Dandekar, S., Van Rompay, K. K. A. and Delwart, E. L. (2004). Number of CD4+ and CD8+ T-cell CDR3 clonotypes expanding during acute infection of macaques with simian immunodeficiency virus. *Virology* **322**, 105-117.

Bevan, M. J. (2004). Helping the CD8+ T-cell response. *Nature Reviews Immunology* **4**, 595-602.

Bill, J., Appel, V. B. and Palmer, E. (1988). An analysis of T-cell receptor variable region gene expression in major histocompatibility complex disparate mice. *Proc Natl Acad Sci U S A* **85**, 9184-8.

Bill, J., Kanagawa, O., Woodland, D. L. and Palmer, E. (1989). The MHC molecule I-E is necessary but not sufficient for the clonal deletion of V beta 11-bearing T-cells. *J Exp Med* **169**, 1405-19.

Bishop, R., Sohanpal, B. K., Allsopp, B. A., Spooner, P. R., Dolan, T. T. and Morzaria, S. P. (1993). Detection of polymorphisms among *Theileria parva* stocks using repetitive, telomeric and ribosomal DNA probes and anti-schizont monoclonal antibodies. *Parasitology* **107**, 19-31.

Bishop, R., Spooner, P. R., Kanhai, G. K., Kiarie, J., Latif, A. A., Hove, T., Masaka, S. and Dolan, T. T. (1994). Molecular characterization of Theileria parasites: application to the epidemiology of theileriosis in Zimbabwe. *Parasitology* **109**, 573-81.

Bishop, R., Spooner, P. R. and Sohanpal, B. K. (1996). Genetic fingerprinting of *Theileria parva* using a telomeric DNA probe. *Parasitology Research* **82**, 264-266.

Bishop, R., Geysen, D., Spooner, P., Skilton, R., Nene, V., Dolan, T. and Morzaria, S. (2001). Molecular and immunological characterisation of *Theileria parva* stocks which are components of the 'Muguga cocktail' used for vaccination against East Coast fever in cattle. *Vet Parasitol* **94**, 227-37.

Bishop, R., Nene, V., Staeyert, J., Rowlands, J., Nyanjui, J., Osaso, J., Morzaria, S. and Musoke, A. (2003). Immunity to East Coast fever in cattle induced by a polypeptide fragment of the major surface coat protein of *Theileria parva* sporozoites. *Vaccine* **21**, 1205-12.

Blackman, M. A., Marrack, P. and Kappler, J. (1989). Influence of the major histocompatibility complex on positive thymic selection of V beta 17a+ T-cells. *Science* **244**, 214-7.

Blattman, J. N., Sourdive, D. J. D., Murali-Krishna, K., Ahmed, R. and Altman, J. D. (2000). Evolution of the T-cell repertoire during primary, memory, and recall responses to viral infection. *J Immunol* **165**, 6081-6090.

Borg, N. A., Ely, L. K., Beddoe, T., Macdonald, W. A., Reid, H. H., Clements, C. S., Purcell, A. W., Kjer-Nielsen, L., Miles, J. J., Burrows, S. R. *et al.* (2005). The CDR3 regions of an immunodominant T-cell receptor dictate the 'energetic landscape' of peptide-MHC recognition. *Nat Immunol* **6**, 171-180.

Borgulya, P., Kishi, H., Uematsu, Y. and von Boehmer, H. (1992). Exclusion and inclusion of alpha and beta T-cell receptor alleles. *Cell* **69**, 529-37.

Bosc, N. and Lefranc, M. P. (2000). The mouse (*Mus musculus*) T-cell receptor beta variable (TRBV), diversity (TRBD) and joining (TRBJ) genes. *Exp Clin Immunogenet* **17**, 216-28.

Bour, H., Michielin, O., Bousso, P., Cerottini, J.-C. and MacDonald, H. R. (1999). Dramatic influence of V-beta gene polymorphism on an antigen-specific CD8+ T-cell response *in vivo*. *J Immunol* **162**, 4647-4656.

Bousso, P., Casrouge, A., Altman, J. D., Haury, M., Kanellopoulos, J., Abastado, J.-P. and Kourilsky, P. (1998). Individual variations in the murine T-cell response to a specific peptide reflect variability in naive repertoires. *Immunity* **9**, 169-178.

Bousso, P., Levraud, J.-P., Kourilsky, P. and Abastado, J.-P. (1999). The composition of a primary T-cell response is largely determined by the timing of recruitment of individual T-cell clones. *J. Exp. Med.* **189**, 1591-1600.

Bouvier, G., Watrin, F., Naspetti, M., Verthuy, C., Naquet, P. and Ferrier, P. (1996). Deletion of the mouse T-cell receptor beta gene enhancer blocks alpha beta T-cell development. *PNAS* **93**, 7877-7881.

Brawand, P., Cerottini, J.-C. and MacDonald, H. R. (1999). Hierarchal utilization of different T-cell receptor V beta gene segments in the CD8+ T-cell response to an immunodominant Moloney Leukemia Virus-encoded epitope *in vivo*. *J. Virol.* **73**, 9161-9169.

Brehm, M. A., Pinto, A. K., Daniels, K. A., Schneck, J. P., Welsh, R. M. and Selin, L. K. (2002). T-cell immunodominance and maintenance of memory regulated by unexpectedly cross-reactive pathogens. *Nat Immunol* **3**, 627-34.

Broeren, C. P. M., van Eden, W., Kusters, J. G., Lenstra, J. A., Georges, M. G. M. and Logtenberg, V. T. (1991). Conserved nucleotide sequences at the 5' end of T-cell receptor variable genes facilitate polymerase chain reaction amplification. *European Journal of Immunology* **21**, 569-575.

Brooks, J. M., Colbert, R. A., Mear, J. P., Leese, A. M. and Rickinson, A. B. (1998). HLA-B27 subtype polymorphism and CTL epitope choice: studies with EBV peptides link immunogenicity with stability of the B27:peptide complex. *J Immunol* **161**, 5252-9.

Brown, W. C., Sugimoto, C. and Grab, D. J. (1989). *Theileria parva*: bovine helper T-cell clones specific for both infected lymphocytes and schizont membrane antigens. *Exp Parasitol* **69**, 234-48.

Buitkamp, J., Schwaiger, F. W. and Epplen, J. T. (1993). VB6 T-cell receptor elements in artiodactyls: conservation and germline polymorphisms. *Mamm Genome* **4**, 504-10.

Bullock, T. N., Colella, T. A. and Engelhard, V. H. (2000). The density of peptides displayed by dendritic cells affects immune responses to human tyrosinase and gp100 in HLA-A2 transgenic mice. *J Immunol* **164**, 2354-61.

Burridge, M. J., Morzaria, S. P., Cunningham, M. P. and Brown, C. G. (1972). Duration of immunity to East Coast fever (*Theileria parva* infection of cattle). *Parasitology* **64**, 511-5.

Burrows, S. R., Silins, S. L., Moss, D. J., Khanna, R., Misko, I. S. and Argaet, V. P. (1995). T-cell receptor repertoire for a viral epitope in humans is diversified by tolerance to a background major histocompatibility complex antigen. *J. Exp. Med.* **182**, 1703-1715.

Busch, D. H. and Pamer, E. G. (1998). MHC class I/peptide stability: implications for immunodominance, *in vitro* proliferation, and diversity of responding CTL. *J Immunol* **160**, 4441-8.

Busch, D. H. and Pamer, E. G. (1999). T-cell affinity maturation by selective expansion during Infection. *J. Exp. Med.* **189**, 701-710.

Butler, J. E., Wertz, N., Sun, J. and Sacco, R. E. (2005). Comparison of the expressed porcine V beta and J beta repertoire of thymocytes and peripheral T-cells. *Immunology* **114**, 184-93.

Butz, E. A. and Bevan, M. J. (1998). Massive expansion of antigen-specific CD8+ T-cells during an acute virus infection. *Immunity* **8**, 167-175.

Caccia, N., Kronenberg, M., Saxe, D., Haars, R., Bruns, G. A., Goverman, J., Malissen, M., Willard, H., Yoshikai, Y., Simon, M. *et al.* (1984). The T-cell receptor beta chain genes are located on chromosome 6 in mice and chromosome 7 in humans. *Cell* **37**, 1091-9.

Callan, M. F. C., Tan, L., Annels, N., Ogg, G. S., Wilson, J. D. K., O'Callaghan, C. A., Steven, N., McMichael, A. J. and Rickinson, A. B. (1998). Direct visualization of antigen-specific CD8+ T-cells during the primary immune response to epstein-barr virus *in vivo*. *J. Exp. Med.* **187**, 1395-1402.

Campos-Lima, P.-O. d., Levitsky, V., Imreh, M. P., Gavioli, R. and Masucci, M. G. (1997). Epitope-dependent selection of highly restricted or diverse T-cell receptor repertoires in response to persistent infection by epstein-barr virus. *J. Exp. Med.* **186**, 83-89.

Cao, W., Myers-Powell, B. A. and Braciale, T. J. (1996). The weak CD8+ CTL response to an influenza hemagglutinin epitope reflects limited T-cell availability. *J Immunol* **157**, 505-11.

Casanova, J. L., Romero, P., Widmann, C., Kourilsky, P. and Maryanski, J. L. (1991). T-cell receptor genes in a series of class I major histocompatibility complex-restricted cytotoxic T lymphocyte clones specific for a *Plasmodium berghei* nonapeptide: implications for T cell allelic exclusion and antigen-specific repertoire. *J. Exp. Med.* **174**, 1371-1383.

Casanova, J. L., Cerottini, J. C., Matthes, M., Necker, A., Gournier, H., Barra, C., Widmann, C., MacDonald, H. R., Lemonnier, F. and Malissen, B. (1992). H-2-restricted cytolytic T lymphocytes specific for HLA display T-cell receptors of limited diversity. *J. Exp. Med.* **176**, 439-447.

Casanova, J. L. and Maryanski, J. L. (1993). Antigen-selected T-cell receptor diversity and self-nonself homology. *Immunol Today* **14**, 391-4.

Casrouge, A., Beaudoing, E., Dalle, S., Pannetier, C., Kanellopoulos, J. and Kourilsky, P. (2000). Size estimate of the alpha-beta TCR repertoire of naive mouse splenocytes. *J Immunol* **164**, 5782-5787.

Cazenave, P. A., Marche, P. N., Jouvin-Marche, E., Voegtle, D., Bonhomme, F., Bandeira, A. and Coutinho, A. (1990). V beta 17 gene polymorphism in wild-derived mouse strains: two amino acid substitutions in the V beta 17 region greatly alter T-cell receptor specificity. *Cell* **63**, 717-28.

Charmley, P., Wang, K., Hood, L. and Nickerson, D. A. (1993). Identification and physical mapping of a polymorphic human T-cell receptor V beta gene with a frequent null allele. *J Exp Med* **177**, 135-43.

Chen, F., Rowen, L., Hood, L. and Rothenberg, E. V. (2001a). Differential transcriptional regulation of individual TCR V beta segments before gene rearrangement. *J Immunol* **166**, 1771-1780.

Chen, W., Khilko, S., Fecondo, J., Margulies, D. H. and McCluskey, J. (1994). Determinant selection of major histocompatibility complex class I-restricted antigenic peptides is explained by class I-peptide affinity and is strongly influenced by nondominant anchor residues. *J Exp Med* **180**, 1471-83.

Chen, W., Anton, L. C., Bennink, J. R. and Yewdell, J. W. (2000). Dissecting the multifactorial causes of immunodominance in class I-restricted T-cell responses to viruses. *Immunity* **12**, 83-93.

Chen, W., Norbury, C. C., Cho, Y., Yewdell, J. W. and Bennink, J. R. (2001b). Immunoproteasomes shape immunodominance hierarchies of antiviral CD8+ T-cells at the levels of T-cell repertoire and presentation of viral antigens. *J Exp Med* **193**, 1319-26.

Chen, Y., Webster, R. G. and Woodland, D. L. (1998). Induction of CD8+ T-cell responses to dominant and subdominant epitopes and protective immunity to Sendai virus infection by DNA vaccination. *J Immunol* **160**, 2425-32.

Chen, Z. W., Li, Y., Zeng, X., Kuroda, M. J., Schmitz, J. E., Shen, Y., Lai, X., Shen, L. and Letvin, N. L. (2001c). The TCR repertoire of an immunodominant CD8+ T lymphocyte population. *J Immunol* **166**, 4525-4533.

Choi, E. Y., Christianson, G. J., Yoshimura, Y., Sproule, T. J., Jung, N., Joyce, S. and Roopenian, D. C. (2002). Immunodominance of H60 is caused by an abnormally high precursor T-cell pool directed against its unique minor histocompatibility antigen peptide. *Immunity* **17**, 593-603.

Chothia, C., Boswell, D. R. and Lesk, A. M. (1988). The outline structure of the T-cell alpha beta receptor. *Embo J* **7**, 3745-55.

Cibotti, R., Cabaniols, J. P., Pannetier, C., Delarbre, C., Vergnon, I., Kanellopoulos, J. M. and Kourilsky, P. (1994). Public and private V beta T-cell receptor repertoires against hen egg white lysozyme (HEL) in nontransgenic versus HEL transgenic mice. *J. Exp. Med.* **180**, 861-872.

Clarencio, J., de Oliveira, C. I., Bomfim, G., Pompeu, M. M., Teixeira, M. J., Barbosa, T. C., Souza-Neto, S., Carvalho, E. M., Brodskyn, C., Barral, A. *et al.* (2006). Characterization of the T-cell receptor V beta repertoire in the human immune response against *Leishmania* parasites. *Infect. Immun.* **74**, 4757-4765.

Clark, S. P., Arden, B., Kabelitz, D. and Mak, T. W. (1995). Comparison of human and mouse T-cell receptor variable gene segment subfamilies. *Immunogenetics* **42**, 531-540.

Cole, G. A., Hogg, T. L., Coppola, M. A. and Woodland, D. L. (1997). Efficient priming of CD8+ memory T-cells specific for a subdominant epitope following Sendai virus infection. *J Immunol* **158**, 4301-4309.

Connolly, J. M. (1994). The peptide p2Ca is immunodominant in allorecognition of Ld by beta chain variable region V beta 8+ but not V beta 8- strains. *Proc Natl Acad Sci U S A* **91**, 11482-6.

Conrad, M. L., Pettman, R., Whitehead, J., McKinnel, L., Davis, S. K. and Koop, B. F. (2002). Genomic sequencing of the bovine T-cell receptor beta locus. *Vet Immunol Immunopathol* **87**, 439-41.

Conrad, P. A., Denham, D. and Brown, C. G. (1986). Intraerythrocytic multiplication of *Theileria parva in vitro*: an ultrastructural study. *Int J Parasitol* **16**, 223-9.

Conrad, P. A., Iams, K., Brown, W. C., Sohanpal, B. and ole-MoiYoi, O. K. (1987a). DNA probes detect genomic diversity in *Theileria parva* stocks. *Mol Biochem Parasitol* **25**, 213-26.

Conrad, P. A., Stagg, D. A., Grootenhuis, J. G., Irvin, A. D., Newson, J., Njamunggeh, R. E., Rossiter, P. B. and Young, A. S. (1987b). Isolation of Theileria parasites from African buffalo (*Syncerus caffer*) and characterization with anti-schizont monoclonal antibodies. *Parasitology* **94**, 413-23.

Cook, J. E. and Beverley, P. C. L. (2001). Analysis of lymphocyte diversity in the elderly: heteroduplex analysis and alternative techniques. *Experimental Gerontology* **36**, 583-589.

Correia-Neves, M., Waltzinger, C., Mathis, D. and Benoist, C. (2001). The shaping of the T-cell repertoire. *Immunity* **14**, 21-32.

Cose, S. C., Kelly, J. M. and Carbone, F. R. (1995). Characterization of diverse primary herpes simplex virus type 1 gB- specific cytotoxic T-cell response showing a preferential V beta bias. *J. Virol.* **69**, 5849-5852.

Coulie, P. G., Karanikas, V., Colau, D., Lurquin, C., Landry, C., Marchand, M., Dorval, T., Brichard, V. and Boon, T. (2001). A monoclonal cytolytic T-lymphocyte response observed in a melanoma patient vaccinated with a tumor-specific antigenic peptide encoded by gene MAGE-3. *PNAS* **98**, 10290-10295.

Craiu, A., Akopian, T., Goldberg, A. and Rock, K. L. (1997). Two distinct proteolytic processes in the generation of a major histocompatibility complex class I-presented peptide. *Proc Natl Acad Sci U S A* **94**, 10850-5.

Creemers, P. (1982). Lack of reactivity of sera from *Theileria parva*-infected and recovered cattle against cell membrane antigens of *Theileria parva* transformed cell lines. *Vet Immunol Immunopathol* **3**, 427-38.

Crotzer, V. L., Christian, R. E., Brooks, J. M., Shabanowitz, J., Settlage, R. E., Marto, J. A., White, F. M., Rickinson, A. B., Hunt, D. F. and Engelhard, V. H. (2000). Immunodominance among EBV-derived epitopes restricted by HLA-B27 does not correlate with epitope abundance in EBV-transformed B-lymphoblastoid cell lines. *J Immunol* **164**, 6120-9.

Cunningham, M. P., Brown, C. G. D., Burridge, M. J. and Purnell, R. E. (1973). Cryopreservation of infective particles of *Theileria parva*. *International Journal for Parasitology* **3**, 583-587.

Daly, K., Nguyen, P., Woodland, D. L. and Blackman, M. A. (1995). Immunodominance of major histocompatibility complex class I-restricted influenza virus epitopes can be influenced by the T-cell receptor repertoire. *J Virol* **69**, 7416-22.

Daubenberger, C. A., Taracha, E. L., Gaidulis, L., Davis, W. C. and McKeever, D. J. (1999). Bovine gamma-delta T-cell responses to the intracellular protozoan parasite *Theileria parva*. *Infect Immun* **67**, 2241-9.

Davies, C. J., Joosten, I., Bernoco, D., Arriens, M. A., Bester, J., Ceriotti, G., Ellis, S., Hensen, E. J., Hines, H. C., Horin, P. *et al.* (1994). Polymorphism of bovine MHC class I genes. Joint report of the Fifth International Bovine Lymphocyte Antigen (BoLA) Workshop, Interlaken, Switzerland, 1 August 1992. *Eur J Immunogenet* **21**, 239-58.

Davis, M. M. and Bjorkman, P. J. (1988). T-cell antigen receptor genes and T-cell recognition. *Nature* **334**, 395-402.

Davis, M. M., Krogsgaard, M., Huppa, J. B., Sumen, C., Purbhoo, M. A., Irvine, D. J., Wu, L. C. and Ehrlich, L. (2003). Dynamics of cell surface molecules during T-cell recognition. *Annual Review of Biochemistry* **72**, 717-742.

Davis, W. C. and Splitter, G. S. (1991). Bovine CD2 (BoCD2). *Veterinary Immunology and Immunopathology* **27**, 43-50.

Davis, W. C., MacHugh, N. D., Park, Y. H., Hamilton, M. J. and Wyatt, C. R. (1993). 6.4 Identification of a monoclonal antibody reactive with the bovine orthologue of CD3 (BoCD3). *Veterinary Immunology and Immunopathology* **39**, 85-91.

Davis, W. C., Brown, W. C., Hamilton, M. J., Wyatt, C. R., Orden, J. A., Khalid, A. M. and Naessens, J. (1996). Analysis of monoclonal antibodies specific for the gamma-delta TCR. *Veterinary Immunology and Immunopathology* **52**, 275-283.

Davodeau, F., Peyrat, M. A., Romagne, F., Necker, A., Hallet, M. M., Vie, H. and Bonneville, M. (1995). Dual T-cell receptor beta chain expression on human T lymphocytes. *J. Exp. Med.* **181**, 1391-1398.

Del Val, M., Schlicht, H. J., Ruppert, T., Reddehase, M. J. and Koszinowski, U. H. (1991). Efficient processing of an antigenic sequence for presentation by MHC class I molecules depends on its neighboring residues in the protein. *Cell* **66**, 1145-53.

DeMartini, J. C. and Moulton, J. E. (1973). Responses of the bovine lymphatic system to infection by *Theileria parva*. I. Histology and ultrastructure of lymph nodes in experimentally-infected calves. *J Comp Pathol* **83**, 281-98.

Dembic, Z., Haas, W., Weiss, S., McCubrey, J., Kiefer, H., von Boehmer, H. and Steinmetz, M. (1986). Transfer of specificity by murine alpha and beta T-cell receptor genes. *Nature* **320**, 232-238.

Deng, Y., Yewdell, J. W., Eisenlohr, L. C. and Bennink, J. R. (1997). MHC affinity, peptide liberation, T cell repertoire, and immunodominance all contribute to the paucity of MHC class I-restricted peptides recognized by antiviral CTL. *J Immunol* **158**, 1507-15.

Dessauge, F., Lizundia, R., Baumgartner, M., Chaussepied, M. and Langsley, G. (2005). Taking the Myc is bad for Theileria. *Trends Parasitol* **21**, 377-85.

Di Meo, G. P., Perucatti, A., Schibler, L., Incarnato, D., Ferrara, L., Cribiu, E. P. and Iannuzzi, L. (2000). Thirteen type I loci from HSA4q, HSA6p, HSA7q and HSA12q were comparatively FISH-mapped in four river buffalo and sheep chromosomes. *Cytogenetic and Genome Research* **90**, 102-105.

Dietrich, P. Y., Walker, P. R., Schnuriger, V., Saas, P., Perrin, G., Guillard, M., Gaudin, C. and Caignard, A. (1997). TCR analysis reveals significant repertoire selection during *in vitro* lymphocyte culture. *Int. Immunol.* **9**, 1073-1083.

Dobbelaere, D. A., Spooner, P. R., Barry, W. C. and Irvin, A. D. (1984). Monoclonal antibody neutralizes the sporozoite stage of different *Theileria parva* stocks. *Parasite Immunol* **6**, 361-70.

Dobbelaere, D. A. and Rottenberg, S. (2003). Theileria-induced leukocyte transformation. *Curr Opin Microbiol* **6**, 377-82.

Dobbelaere, D. A. and Kuenzi, P. (2004). The strategies of the Theileria parasite: a new twist in host-pathogen interactions. *Curr Opin Immunol* **16**, 524-30.

Dolan, T. T., Radley, D. E., Brown, C. G. D., Cunningham, M. P., Morzaria, S. P. and Young, A. S. (1980). East coast fever: 4. Further studies on the protection of cattle immunized with a combination of theilerial strains. *Veterinary Parasitology* **6**, 325-332.

Douek, D. C., Betts, M. R., Brenchley, J. M., Hill, B. J., Ambrozak, D. R., Ngai, K.-L., Karandikar, N. J., Casazza, J. P. and Koup, R. A. (2002). A novel approach to the analysis of specificity, clonality, and frequency of HIV-specific T-cell responses reveals a potential mechanism for control of viral escape. *J Immunol* **168**, 3099-3104.

Eggers, M., Boes-Fabian, B., Ruppert, T., Kloetzel, P. M. and Koszinowski, U. H. (1995). The cleavage preference of the proteasome governs the yield of antigenic peptides. *J Exp Med* **182**, 1865-70.

Eisenlohr, L. C., Yewdell, J. W. and Bennink, J. R. (1992). Flanking sequences influence the presentation of an endogenously synthesized peptide to cytotoxic T lymphocytes. *J Exp Med* **175**, 481-7.

Eleuteri, A. M., Kohanski, R. A., Cardozo, C. and Orlowski, M. (1997). Bovine spleen multicatalytic proteinase complex (proteasome). Replacement of X, Y, and Z subunits by LMP7, LMP2, and MECL1 and changes in properties and specificity. *J Biol Chem* **272**, 11824-31.

Ellis, J. A., Davis, W. C., MacHugh, N. D., Emery, D. L., Kaushal, A. and Morrison, W. I. (1988). Differentiation antigens on bovine mononuclear phagocytes identified by monoclonal antibodies. *Veterinary Immunology and Immunopathology* **19**, 325-340.

Ellis, S. A., Holmes, E. C., Staines, K. A., Smith, K. B., Stear, M. J., McKeever, D. J., MacHugh, N. D. and Morrison, W. I. (1999). Variation in the number of expressed MHC genes in different cattle class I haplotypes. *Immunogenetics* **50**, 319-28.

Ema, H., Cumano, A. and Kourilsky, P. (1997). TCR-beta repertoire development in the mouse embryo. *J Immunol* **159**, 4227-4232.

Emery, D. L. (1981a). Adoptive transfer of immunity to infection with *Theileria parva* (East Coast fever) between cattle twins. *Res Vet Sci* **30**, 364-7.

Emery, D. L. (1981b). Kinetics of infection with *Theileria parva* (East Coast fever) in the central lymph of cattle. *Vet Parasitol* **9**, 1-16.

Emery, D. L., Eugui, E. M., Nelson, R. T. and Tenywa, T. (1981). Cell-mediated immune responses to *Theileria parva* (East Coast fever) during immunization and lethal infections in cattle. *Immunology* **43**, 323-36.

Emery, D. L., Morrison, W. I., Buscher, G. and Nelson, R. T. (1982). Generation of cell-mediated cytotoxicity to *Theileria parva* (East Coast fever) after inoculation of cattle with parasitized lymphoblasts. *J Immunol* **128**, 195-200.

Eugui, E. M. and Emery, D. L. (1981). Genetically restricted cell-mediated cytotoxicity in cattle immune to *Theileria parva*. *Nature* **290**, 251-4.

Fawcett, D. W., Buscher, G. and Doxsey, S. (1982a). Salivary gland of the tick vector of East Coast fever. III. The ultrastructure of sporogony in *Theileria parva*. *Tissue Cell* **14**, 183-206.

Fawcett, D. W., Buscher, G. and Doxsey, S. (1982b). Salivary gland of the tick vector of East Coast fever. IV. Cell type selectivity and host cell responses to *Theileria parva*. *Tissue Cell* **14**, 397-414.

Fawcett, D. W., Doxsey, S., Stagg, D. A. and Young, A. S. (1982c). The entry of sporozoites of *Theileria parva* into bovine lymphocytes *in vitro*. Electron microscopic observations. *Eur J Cell Biol* **27**, 10-21.

Fawcett, D. W., Conrad, P. A., Grootenhuis, J. G. and Morzaria, S. P. (1987). Ultrastructure of the intra-erythrocytic stage of Theileria species from cattle and waterbuck. *Tissue Cell* **19**, 643-55.

Feeney, A. J., Goebel, P. and Espinoza, C. R. (2004). Many levels of control of V gene rearrangement frequency. *Immunological Reviews* **200**, 44-56.

Folch, G. and Lefranc, M. P. (2000a). The human T-cell receptor beta diversity (TRBD) and beta joining (TRBJ) genes. *Exp Clin Immunogenet* **17**, 107-14.

Folch, G. and Lefranc, M. P. (2000b). The human T-cell receptor beta variable (TRBV) genes. *Exp Clin Immunogenet* **17**, 42-54.

Frahm, N., Kiepiela, P., Adams, S., Linde, C. H., Hewitt, H. S., Sango, K., Feeney, M. E., Addo, M. M., Lichterfeld, M., Lahaie, M. P. *et al.* (2006). Control of human immunodeficiency virus replication by cytotoxic T lymphocytes targeting subdominant epitopes. *Nat Immunol* **7**, 173-8.

Fu, T. M., Friedman, A., Ulmer, J. B., Liu, M. A. and Donnelly, J. J. (1997). Protective cellular immunity: cytotoxic T-lymphocyte responses against dominant and recessive epitopes of influenza virus nucleoprotein induced by DNA immunization. *J Virol* **71**, 2715-21.

Funkhouser, W., Koop, B. F., Charmley, P., Martindale, D., Slightom, J. and Hood, L. (1997). Evolution and selection of primate T-cell antigen receptor BV8 gene subfamily. *Molecular Phylogenetics and Evolution* **8**, 51-64.

Gahm, S., Fowlkes, B. J., Jameson, S. C., Gascoigne, N. R. J., Cotterman, M. M., Kanagawa, O., Schwartz, R. H. and Matis, L. A. (1991). Profound alteration in an alpha-beta T-cell antigen receptor repertoire due to polymorphism in the first complementarity-determining region of the beta chain. *PNAS* **88**, 10267-10271.

Gallard, A., Foucras, G., Coureau, C. and Guery, J.-C. (2002). Tracking T cell clonotypes in complex T lymphocyte populations by real-time quantitative PCR using fluorogenic complementarity-determining region-3-specific probes. *Journal of Immunological Methods* **270**, 269-280.

Gallimore, A., Hombach, J., Dumrese, T., Rammensee, H. G., Zinkernagel, R. M. and Hengartner, H. (1998). A protective cytotoxic T-cell response to a subdominant epitope is influenced by the stability of the MHC class I/peptide complex and the overall spectrum of viral peptides generated within infected cells. *Eur J Immunol* **28**, 3301-11.

Garboczi, D. N., Ghosh, P., Utz, U., Fan, Q. R., Biddison, W. E. and Wiley, D. C. (1996). Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* **384**, 134-141.

Garcia, K. C., Degano, M., Pease, L. R., Huang, M., Peterson, P. A., Teyton, L. and Wilson, I. A. (1998). Structural basis of plasticity in T-cell receptor recognition of a self peptide-MHC antigen. *Science* **279**, 1166-1172.

Garcia, K. C., Degano, M., Stanfield, R. L., Brunmark, A., Jackson, M. R., Peterson, P. A., Teyton, L. and Wilson, I. A. (1996). An alpha beta T-cell receptor structure at 2.5 A and its orientation in the TCR-MHC complex. *Science* **274**, 209-219.

Garcia, K. C. and Adams, E. J. (2005). How the T-cell receptor sees antigen - a structural view. *Cell* **122**, 333-336.

Gardner, M. J., Bishop, R., Shah, T., de Villiers, E. P., Carlton, J. M., Hall, N., Ren, Q., Paulsen, I. T., Pain, A., Berriman, M. *et al.* (2005). Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science* **309**, 134-137.

Geysen, D., Bishop, R., Skilton, R., Dolan, T. T. and Morzaria, S. (1999). Molecular epidemiology of *Theileria parva* in the field. *Trop Med Int Health* **4**, A21-7.

Gileadi, U., Moins-Teisserenc, H. T., Correa, I., Booth, B. L., Jr., Dunbar, P. R., Sewell, A. K., Trowsdale, J., Phillips, R. E. and Cerundolo, V. (1999). Generation of an immunodominant CTL epitope is affected by proteasome subunit composition and stability of the antigenic protein. *J Immunol* **163**, 6045-52.

Gillespie, G. M. A., Wills, M. R., Appay, V., O'Callaghan, C., Murphy, M., Smith, N., Sissons, P., Rowland-Jones, S., Bell, J. I. and Moss, P. A. H. (2000). Functional heterogeneity and high frequencies of cytomegalovirus-Specific CD8+ T lymphocytes in healthy seropositive donors. *J. Virol.* **74**, 8140-8150.

Gillespie, G. M. A., Stewart-Jones, G., Rengasamy, J., Beattie, T., Bwayo, J. J., Plummer, F. A., Kaul, R., McMichael, A. J., Easterbrook, P., Dong, T. *et al.* (2006). Strong TCR conservation and altered T-cell cross-reactivity characterize a B*57-restricted immune response in HIV-1 infection. *J Immunol* **177**, 3893-3902.

Giudicelli, V. and Lefranc, M.-P. (1999). Ontology for immunogenetics: the IMGT-ONTOLOGY. *Bioinformatics* **15**, 1047-1054.

Giudicelli, V., Chaume, D. and Lefranc, M. P. (2004). IMGT/V-QUEST, an integrated software program for immunoglobulin and T-cell receptor V-J and V-D-J rearrangement analysis. *Nucleic Acids Res* **32**, W435-40.

Glusman, G., Rowen, L., Lee, I., Boysen, C., Roach, J. C., Smit, A. F., Wang, K., Koop, B. F. and Hood, L. (2001). Comparative genomics of the human and mouse T cell receptor loci. *Immunity* **15**, 337-49.

Goddeeris, B. M., Morrison, W. I. and Teale, A. J. (1986). Generation of bovine cytotoxic cell lines, specific for cells infected with the protozoan parasite *Theileria parva* and restricted by products of the major histocompatibility complex. *Eur J Immunol* **16**, 1243-9.

Goddeeris, B. M. and Morrison, W. I. (1988). Tehcniques for the generation, cloning, and characterisation of bovine cytotoxic T-cells specific for the protozoan *Theileria parva*. *Journal of Tissue Culture Methods* **11**, 101-110.

Goddeeris, B. M., Morrison, W. I., Toye, P. G. and Bishop, R. (1990). Strain specificity of bovine *Theileria parva*-specific cytotoxic T cells is determined by the phenotype of the restricting class I MHC. *Immunology* **69**, 39-44.

Goddeeris, B. M., Dunlap, S., Bensaid, A., MacHugh, N. D. and Morrison, W. I. (1991). Cell surface phenotype of two cloned populations of bovine lymphocytes displaying non-specific cytotoxic activity. *Veterinary Immunology and Immunopathology* **27**, 195-199.

Goldrath, A. W. and Bevan, M. J. (1999). Selecting and maintaining a diverse T-cell repertoire. *Nature* **402**, 255-62.

Goulder, P. J., Sewell, A. K., Lalloo, D. G., Price, D. A., Whelan, J. A., Evans, J., Taylor, G. P., Luzzi, G., Giangrande, P., Phillips, R. E. ***et al.*** (1997). Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte responses in two human histocompatibility leukocyte antigens (HLA)-identical siblings with HLA-A*0201 are influenced by epitope mutation. *J Exp Med* **185**, 1423-33.

Graham, S. P., Pelle, R., Honda, Y., Mwangi, D. M., Tonukari, N. J., Yamage, M., Glew, E. J., de Villiers, E. P., Shah, T., Bishop, R. ***et al.*** (2006). *Theileria parva* candidate vaccine antigens recognized by immune bovine cytotoxic T lymphocytes. *PNAS* **103**, 3286-3291.

Groettrup, M., Ruppert, T., Kuehn, L., Seeger, M., Standera, S., Koszinowski, U. and Kloetzel, P. M. (1995). The interferon-γ-inducible 11 S regulator (PA28) and the LMP2/LMP7 subunits govern the peptide production by the 20 S proteasome *in vitro*. *J. Biol. Chem.* **270**, 23808-23815.

Grufman, P., Sandberg, J. K., Wolpert, E. Z. and Karre, K. (1999). Immunization with dendritic cells breaks immunodominance in CTL responses against minor histocompatibility and synthetic peptide antigens. *J Leukoc Biol* **66**, 268-71.

Gulwani-Akolkar, B., Posnett, D. N., Janson, C. H., Grunewald, J., Wigzell, H., Akolkar, P., Gregersen, P. K. and Silver, J. (1991). T-cell receptor V-segment frequencies in peripheral blood T cells correlate with human leukocyte antigen type. *J Exp Med* **174**, 1139-46.

Halle, J. P., Haus-Seuffert, P., Woltering, C., Stelzer, G. and Meisterernst, M. (1997). A conserved tissue-specific structure at a human T-cell receptor beta- chain core promoter. *Mol. Cell. Biol.* **17**, 4220-4229.

Halsey, W. A., Jr., Palmer, B. E., DeMartini, J. C. and Howell, M. D. (1999). Analysis of sheep T-cell receptor beta-chain heterogeneity. *Immunogenetics* **49**, 206-14.

Hamrouni, A., Aublin, A., Guillaume, P. and Maryanski, J. L. (2003). T-cell receptor gene rearrangement lineage analysis reveals clues for the origin of Highly restricted antigen-specific repertoires. *J. Exp. Med.* **197**, 601-614.

Haqqi, T. M., Banerjee, S., Anderson, G. D. and David, C. S. (1989). RIII S/J (H-2r). An inbred mouse strain with a massive deletion of T-cell receptor V beta genes. *J Exp Med* **169**, 1903-9.

Hassanin, A. and Douzery, E. J. (2003). Molecular and morphological phylogenies of ruminantia and the alternative position of the moschidae. *Syst Biol* **52**, 206-28.

Hawes, G. E., Struyk, L. and van den Elsen, P. J. (1993). Differential usage of T-cell receptor V gene segments in CD4+ and CD8+ subsets of T lymphocytes in monozygotic twins. *J Immunol* **150**, 2033-45.

Heath, W. R. and Miller, J. F. (1993). Expression of two alpha chains on the surface of T-cells in T-cell receptor transgenic mice. *J. Exp. Med.* **178**, 1807-1811.

Hennecke, J. and Wiley, D. C. (2001). T-cell receptor-MHC interactions up close. *Cell* **104**, 1-4.

Henry, C., Marbrook, J., Vann, D. C. *et al.* (1980). Limiting Dilution Analysis. San Fransico: W.H. Freeman and Co.

Herman, A., Kappler, J. W., Marrack, P. and Pullen, A. M. (1991). Superantigens: mechanism of T-cell stimulation and role in immune responses. *Annual Review of Immunology* **9**, 745-772.

Herzig, C., Blumerman, S., Lefranc, M. P. and Baldwin, C. (2006a). Bovine T-cell receptor gamma variable and constant genes: combinatorial usage by circulating gammadelta T cells. *Immunogenetics* **58**, 138-51.

Herzig, C. T., Blumerman, S. L. and Baldwin, C. L. (2006b). Identification of three new bovine T-cell receptor delta variable gene subgroups expressed by peripheral blood T-cells. *Immunogenetics* **58**, 746-57.

Hesse, J. E., Lieber, M. R., Mizuuchi, K. and Gellert, M. (1989). V(D)J recombination: a functional definition of the joining signals. *Genes Dev.* **3**, 1053-1061.

Heussler, V. T., Taracha, E. L., Musoke, A., Duchateau, L. and McKeever, D. J. (1998). Immunisation with live attenuated *Salmonella dublin* expressing a sporozoite protein confers partial protection against *Theileria parva*. *Vaccine* **16**, 834-41.

Hohlfeld, R. (2003). Hunting (auto)immune T-cells in neuroimmunological diseases. *Brain* **126**, 2-4.

Horwitz, M. S., Yanagi, Y. and Oldstone, M. B. (1994). T-cell receptors from virus-specific cytotoxic T lymphocytes recognizing a single immunodominant nine-amino-acid viral epitope show marked diversity. *J. Virol.* **68**, 352-357.

Housset, D., Mazza, G., Gregoire, C., Piras, C., Malissen, B. and Fontecilla-Camps, J. C. (1997). The three-dimensional structure of a T-cell antigen receptor V alpha V beta heterodimer reveals a novel arrangement of the V beta domain. *Embo J* **16**, 4205-16.

Housset, D. and Malissen, B. (2003). What do TCR-pMHC crystal structures teach us about MHC restriction and alloreactivity? *Trends in Immunology* **24**, 429-437.

Houston, E. F. (1997). The Potential Role Of Superantigens In The Pathogenesis Of Bovine Theileriosis. *PhD Thesis* **University of Glsgow**.

Houston, E. F. and Morrison, W. I. (1999). Identification of seven new TCRBV subfamilies in cattle (Bos taurus). *Eur J Immunogenet* **26**, 349-53.

Houston, E. F., Connelley, T., Parsons, K., MacHugh, N. D. and Morrison, W. I. (2005). Analysis of T-cell receptor BV gene sequences in cattle reveals extensive duplication within the BV9 and BV20 subgroups. *Immunogenetics* **57**, 674-81.

Hughes, M. M., Tillman, R. E., Wehrly, T. D., White, J. M. and Sleckman, B. P. (2003). The B12/23 restriction is critically dependent on recombination signal nonamer and spacer sequences. *J Immunol* **171**, 6604-10.

Hulliger, L., Wilde, K. H., Brown, C. G. and Turner, L. (1964). Mode of multiplication of Theileria in cultures of bovine lymphocytic cells. *Nature* **203**, 728-30.

Irvin, A. D., Ocama, J. G. and Spooner, P. R. (1982). Cycle of bovine lymphoblastoid cells parasitised by *Theileria parva*. *Res Vet Sci* **33**, 298-304.

Irvin, A. D., Dobbelaere, D. A., Mwamachi, D. M., Minami, T., Spooner, P. R. and Ocama, J. G. (1983). Immunisation against East Coast fever: correlation between monoclonal antibody profiles of *Theileria parva* stocks and cross immunity *in vivo*. *Res Vet Sci* **35**, 341-6.

Irvin, A. D. and Mwamachi, D. M. (1983). Clinical and diagnostic features of East Coast fever (*Theileria parva*) infection of cattle. *Vet Rec* **113**, 192-8.

Ishikawa, T., Kono, D., Chung, J., Fowler, P., Theofilopoulos, A., Kakumu, S. and Chisari, F. V. (1998). Polyclonality and multispecificity of the CTL response to a single viral epitope. *J Immunol* **161**, 5842-5850.

Isobe, M., Erikson, J., Emanuel, B. S., Nowell, P. C. and Croce, C. M. (1985). Location of gene for beta subunit of human T-cell receptor at band 7q35, a region prone to rearrangements in T cells. *Science* **228**, 580-2.

Isono, T., Isegawa, Y. and Seto, A. (1994). Sequence and diversity of variable gene segments coding for rabbit T-cell receptor beta chains. *Immunogenetics* **39**, 243-8.

Ito, K., Tsunoda, M., Watanabe, K., Ito, K., Kashiwagi, N. and Obata, F. (1993). Isolation and sequence analysis of cDNA for the dog T-cell receptor TCR alpha and TCR beta chains. *Immunogenetics* **38**, 60-3.

Jores, R. and Meo, T. (1993). Few V gene segments dominate the T-cell receptor beta-chain repertoire of the human thymus. *J Immunol* **151**, 6110-6122.

Jouvin-Marche, E., Trede, N. S., Bandeira, A., Tomas, A., Loh, D. Y. and Cazenave, P. A. (1989). Different large deletions of T-cell receptor V beta genes in natural populations of mice. *Eur J Immunol* **19**, 1921-6.

Jung, D., Bassing, C. H., Fugmann, S. D., Cheng, H. L., Schatz, D. G. and Alt, F. W. (2003). Extrachromosomal recombination substrates recapitulate beyond 12/23 restricted VDJ recombination in nonlymphoid cells. *Immunity* **18**, 65-74.

Kaba, S. A., Musoke, A. J., Schaap, D., Schetters, T., Rowlands, J., Vermeulen, A. N., Nene, V., Vlak, J. M. and van Oers, M. M. (2005). Novel baculovirus-derived p67 subunit vaccines efficacious against East Coast fever in cattle. *Vaccine* **23**, 2791-800.

Kalergis, A. M., Boucheron, N., Doucey, M. A., Palmieri, E., Goyarts, E. C., Vegh, Z., Luescher, I. F. and Nathenson, S. G. (2001). Efficient T-cell activation requires an optimal dwell-time of interaction between the TCR and the pMHC complex. *Nat Immunol* **2**, 229-34.

Kappler, J. W., Roehm, N. and Marrack, P. (1987). T-cell tolerance by clonal elimination in the thymus. *Cell* **49**, 273-80.

Kappler, J. W., Staerz, U., White, J. and Marrack, P. C. (1988). Self-tolerance eliminates T-cells specific for Mls-modified products of the major histocompatibility complex. *Nature* **332**, 35-40.

Kariuki, D. P., Young, A. S., Morzaria, S. P., Lesan, A. C., Mining, S. K., Omwoyo, P., Wafula, J. L. and Molyneux, D. H. (1995). *Theileria parva* carrier state in naturally infected and artificially immunised cattle. *Trop Anim Health Prod* **27**, 15-25.

Kay, R. A., Snowden, N., Hajeer, A. H., Boylston, A. W. and Ollier, W. E. (1994). Genetic control of the human V beta 13.2 T cell repertoire: importance of allelic variation outside the coding regions of the TCRBV13S2 gene. *Eur J Immunol* **24**, 2863-7.

Kazatchkine, M. D. (1995). Nomenclature for T-cell receptor (TCR) gene segments of the immune system. *Immunogenetics* **42**, 451-453.

Kedl, R. M., Rees, W. A., Hildeman, D. A., Schaefer, B., Mitchell, T., Kappler, J. and Marrack, P. (2000). T-cells compete for access to antigen-bearing antigen-presenting cells. *J Exp Med* **192**, 1105-13.

Kedl, R. M., Schaefer, B. C., Kappler, J. W. and Marrack, P. (2002). T-cells down-modulate peptide-MHC complexes on APCs *in vivo*. *Nat Immunol* **3**, 27-32.

Kedl, R. M., Kappler, J. W. and Marrack, P. (2003). Epitope dominance, competition and T-cell affinity maturation. *Curr Opin Immunol* **15**, 120-7.

Kedzierska, K., Turner, S. J. and Doherty, P. C. (2004). Conserved T-cell receptor usage in primary and recall responses to an immunodominant influenza virus nucleoprotein epitope. *PNAS* **101**, 4942-4947.

Khor, B. and Sleckman, B. P. (2002). Allelic exclusion at the TCRβ locus. *Current Opinion in Immunology* **14**, 230-234.

Kim, M. and Kim, K. (2002). Diversity and complexity of CD8+ T-cell responses against a single epitope of adenovirus E1B. *Virology* **295**, 238-249.

Kjer-Nielsen, L., Clements, C. S., Brooks, A. G., Purcell, A. W., McCluskey, J. and Rossjohn, J. (2002). The 1.5 A crystal structure of a highly selected antiviral T-cell receptor provides evidence for a structural basis of immunodominance. *Structure (Camb)* **10**, 1521-32.

Kjer-Nielsen, L., Clements, C. S., Purcell, A. W., Brooks, A. G., Whisstock, J. C., Burrows, S. R., McCluskey, J. and Rossjohn, J. (2003). A structural basis for the selection of dominant αβ T-cell receptors in antiviral immunity. *Immunity* **18**, 53-64.

Komatsu, M., Lamoyi, E. and Mage, R. G. (1987). Genomic DNA encoding rabbit T cell receptor beta-chains: isotypes and allotypes of C beta. *J Immunol* **138**, 1621-1626.

Koop, B. F., Rowen, L., Wang, K., Kuo, C. L., Seto, D., Lenstra, J. A., Howard, S., Shan, W., Deshpande, P. and Hood, L. (1994). The human T-cell receptor TCRAC/TCRDC region: Organization, sequence, and evolution of 97.6 kb of DNA. *Genomics* **19**, 478-493.

Korbel, D. S., Finney, O. C. and Riley, E. M. (2004). Natural killer cells and innate immunity to protozoan pathogens. *International Journal for Parasitology* **34**, 1517-1528.

Kotzin, B. L., Barr, V. L. and Palmer, E. (1985). A large deletion within the T-cell receptor beta-chain gene complex in New Zealand white mice. *Science* **229**, 167-71.

Krimpenfort, P., de Jong, R., Uematsu, Y., Dembic, Z., Ryser, S., von Boehmer, H., Steinmetz, M. and Berns, A. (1988). Transcription of T-cell receptor beta-chain genes is controlled by a downstream regulatory element. *Embo J* **7**, 745-50.

Kumar, S. and Hedges, S. B. (1998). A molecular timescale for vertebrate evolution. *Nature* **392**, 917-920.

Kumar, V. and Sercarz, E. (1994). Holes in the T-cell repertoire to myelin basic protein owing to the absence of the D beta 2-J beta 2 gene cluster: implications for T-cell receptor recognition and autoimmunity. *J. Exp. Med.* **179**, 1637-1643.

La Gruta, N. L., Kedzierska, K., Pang, K., Webby, R., Davenport, M., Chen, W., Turner, S. J. and Doherty, P. C. (2006). A virus-specific CD8+ T-cell immunodominance hierarchy determined by antigen dose and precursor frequencies. *PNAS* **103**, 994-999.

Lacey, S. F., Villacres, M. C., La Rosa, C., Wang, Z., Longmate, J., Martinez, J., Brewer, J. C., Mekhoubad, S., Maas, R., Leedom, J. M. *et al.* (2003). Relative dominance of HLA-B*07 restricted CD8+ T-lymphocyte immune responses to human cytomegalovirus pp65 in persons sharing HLA-A*02 and HLA-B*07 alleles. *Hum Immunol* **64**, 440-52.

Lai, E., Concannon, P. and Hood, L. (1988). Conserved organization of the human and murine T-cell receptor beta-gene families. *Nature* **331**, 543-6.

Lawrence, J. A., Norval, R. A. and Uilenberg, G. (1983). Rhipicephalus zambeziensis as a vector of bovine Theileriae. *Trop Anim Health Prod* **15**, 39-42.

Leduc, I., Hempel, W. M., Mathieu, N., Verthuy, C., Bouvier, G., Watrin, F. and Ferrier, P. (2000). T-cell development in TCRβ enhancer-deleted mice: implications for αβT-cell lineage commitment and differentiation. *J Immunol* **165**, 1364-1373.

Lefranc, M. P., Pommie, C., Ruiz, M., Giudicelli, V., Foulquier, E., Truong, L., Thouvenin-Contet, V. and Lefranc, G. (2003). IMGT unique numbering for immunoglobulin and T-cell receptor variable domains and Ig superfamily V-like domains. *Dev Comp Immunol* **27**, 55-77.

Lehner, P. J., Wang, E. C., Moss, P. A., Williams, S., Platt, K., Friedman, S. M., Bell, J. I. and Borysiewicz, L. K. (1995). Human HLA-A0201-restricted cytotoxic T lymphocyte recognition of influenza A is dominated by T-cells bearing the V beta 17 gene segment. *J. Exp. Med.* **181**, 79-91.

Lessard, P., L'Eplattenier, R., Norval, R. A., Kundert, K., Dolan, T. T., Croze, H., Walker, J. B., Irvin, A. D. and Perry, B. D. (1990). Geographical information systems for studying the epidemiology of cattle diseases caused by *Theileria parva*. *Vet Rec* **126**, 255-62.

Levitsky, V., Zhang, Q. J., Levitskaya, J. and Masucci, M. G. (1996). The life span of major histocompatibility complex-peptide complexes influences the efficiency of presentation and immunogenicity of two class I-restricted cytotoxic T lymphocyte epitopes in the Epstein-Barr virus nuclear antigen 4. *J Exp Med* **183**, 915-26.

Levitsky, V., de Campos-Lima, P.-O., Frisan, T. and Masucci, M. G. (1998). The clonal composition of a peptide-specific oligoclonal CTL repertoire selected in response to persistent EBV infection is stable over time. *J Immunol* **161**, 594-601.

Levraud, J. P., Pannetier, C., Langlade-Demoyen, P., Brichard, V. and Kourilsky, P. (1996). Recurrent T-cell receptor rearrangements in the cytotoxic T lymphocyte response *in vivo* against the p815 murine tumor. *J. Exp. Med.* **183**, 439-449.

Li, W., Gouy, M., Sharp, P. M., O'Huigin, C. and Yang, Y. (1990). Molecular phylogeny of Rodentia, Lagomorpha, Primates, Artiodactyla, and Carnivora and molecular clocks. *PNAS* **87**, 6703-6707.

Liao, N. S., Maltzman, J. and Raulet, D. H. (1989). Positive selection determines T-cell receptor V beta 14 gene usage by CD8+ T cells. *J Exp Med* **170**, 135-43.

Liao, N. S., Maltzman, J. and Raulet, D. H. (1990). Expression of the V beta 5.1 gene by murine peripheral T cells is controlled by MHC genes and skewed to the CD8+ subset. *J Immunol* **144**, 844-8.

Listman, J. A., Rimm, I. J., Wang, Y., Geller, M. C., Tang, J. C., Ho, S., Finn, P. W. and Perkins, D. L. (1996). Plasticity of the T-cell receptor repertoire in TCR beta-hain transgenic mice. *Cellular Immunology* **167**, 44-55.

Liu, J., Ewald, B. A., Lynch, D. M., Nanda, A., Sumida, S. M. and Barouch, D. H. (2006). Modulation of DNA vaccine-elicited CD8+ T lymphocyte epitope immunodominance hierarchies. *J. Virol.*, (epub ahead of publication).

Livak, F., Burtrum, D. B., Rowen, L., Schatz, D. G. and Petrie, H. T. (2000). Genetic modulation of T-cell receptor gene segment usage during somatic recombination. *J. Exp. Med.* **192**, 1191-1196.

Livak, F. (2003). Evolutionarily conserved pattern of gene segment usage within the mammalian TCRÃ locus. *Immunogenetics* **55**, 307-314.

Livak, F. (2004). *In vitro* and *in vivo* studies on the generation of the primary T-cell receptor repertoire. *Immunological Reviews* **200**, 23-35.

Lodoen, M. B. and Lanier, L. L. (2006). Natural killer cells as an initial defense against pathogens. *Current Opinion in Immunology* **18**, 391-398.

Loh, E. Y., Elliott, J. F., Cwirla, S., Lanier, L. L. and Davis, M. M. (1989). Polymerase chain reaction with single-sided specificity: analysis of T-cell receptor delta chain. *Science* **243**, 217-220.

Lopes, A. R., Jaye, A., Dorrell, L., Sabally, S., Alabi, A., Jones, N. A., Flower, D. R., De Groot, A., Newton, P., Lascar, R. M. *et al.* (2003). Greater CD8+ TCR heterogeneity and functional flexibility in HIV-2 compared to HIV-1 infection. *J Immunol* **171**, 307-316.

MacDonald, H. R., Lees, R. K., Schneider, R., Zinkernagel, R. M. and Hengartner, H. (1988a). Positive selection of CD4+ thymocytes controlled by MHC class II gene products. *Nature* **336**, 471-3.

MacDonald, H. R., Schneider, R., Lees, R. K., Howe, R. C., Acha-Orbea, H., Festenstein, H., Zinkernagel, R. M. and Hengartner, H. (1988b). T-cell receptor V beta use predicts reactivity and tolerance to Mlsa-encoded antigens. *Nature* **332**, 40-5.

MacDonald, H. R., Casanova, J. L., Maryanski, J. L. and Cerottini, J. C. (1993). Oligoclonal expansion of major histocompatibility complex class I-restricted cytolytic T lymphocytes during a primary immune response *in vivo*: direct monitoring by flow cytometry and polymerase chain reaction. *J Exp Med* **177**, 1487-92.

MacHugh, N. D. and Sopp, P. (1991). Bovine CD8 (BoCD8). *Veterinary Immunology and Immunopathology* **27**, 65-69.

Maini, Wedderburn, Hall, Wack, Casorati and Beverley, P. (1998). A comparison of two techniques for the molecular tracking of specific T-cell responses; CD4+ human T-cell clones persist in a stable hierarchy but at a lower frequency than clones in the CD8+ population. *Immunology* **94**, 529-535.

Maini, M. K., Gudgeon, N., Wedderburn, L. R., Rickinson, A. B. and Beverley, P. C. L. (2000). Clonal expansions in acute EBV infection are detectable in the CD8+ and not the CD4+ subset and persist with a variable CD45 phenotype. *J Immunol* **165**, 5729-5737.

Manfras, B. J., Terjung, D. and Boehm, B. O. (1999). Non-productive human TCR beta-chain genes represent V-D-J diversity before selection upon function: insight into biased usage of TCRBD and TCRBJ genes and diversity of CDR3 region length. *Human Immunology* **60**, 1090-1100.

Manning, T. C., Schlueter, C. J., Brodnicki, T. C., Parke, E. A., Speir, J. A., Garcia, K. C., Teyton, L., Wilson, I. A. and Kranz, D. M. (1998). Alanine scanning mutagenesis of an αβ T-cell receptor: mapping the energy of antigen recognition. *Immunity* **8**, 413-425.

Marcotty, T., Brandt, J., Billiouw, M., Chaka, G., Losson, B. and Berkvens, D. (2002). Immunisation against *Theileria parva* in eastern Zambia: influence of maternal antibodies and demonstration of the carrier status. *Veterinary Parasitology* **110**, 45-56.

Maritim, A. C., Young, A. S., Lesan, A. C., Ndungu, S. G., Mutugi, J. J. and Stagg, D. A. (1989). Theilerial parasites isolated from carrier cattle after immunization with *Theileria parva* by the infection and treatment method. *Parasitology* **99**, 139-47.

Marrack, P., Bender, J., Jordan, M., Rees, W., Robertson, J., Schaefer, B. C. and Kappler, J. (2001). Major histocompatibility complex proteins and TCRs: do they really go together like a horse and carriage? *J Immunol* **167**, 617-21.

Maryanski, J. L., Jongeneel, C. V., Bucher, P., Casanova, J. L. and Walker, P. R. (1996). Single-cell PCR analysis of TCR repertoires selected by antigen *in vivo*: a high magnitude CD8+ response is comprised of very few clones. *Immunity* **4**, 47-55.

Maryanski, J. L., Casanova, J. L., Falk, K., Gournier, H., Jaulin, C., Kourilsky, P., Lemonnier, F. A., Luthy, R., Rammensee, H. G., Rotzschke, O. *et al.* (1997). The diversity of antigen-specific TCR repertoires reflects the relative complexity of epitopes recognized. *Hum Immunol* **54**, 117-28.

Maryanski, J. L., Attuil, V., Bucher, P. and Walker, P. R. (1999). A quantitative, single-cell PCR analysis of an antigen-specific TCR repertoire selected during an *in vivo* CD8+ response: direct evidence for a wide range of clone sizes with uniform tissue distribution. *Mol Immunol* **36**, 745-53.

Maryanski, J. L., Attuil, V., Hamrouni, A., Mutin, M., Rossi, M., Aublin, A. and Bucher, P. (2001). Individuality of Ag-selected and preimmune TCR repertoires. *Immunol Res* **23**, 75-84.

Mathieu, N., Hempel, W. M., Spicuglia, S., Verthuy, C. and Ferrier, P. (2000). Chromatin remodeling by the TCRβ gene enhancer during early T-cell development: implications for the control of TCRβ locus recombination. *J. Exp. Med.* **192**, 625-636.

Max, E. E., Seidman, J. G. and Leder, P. (1979). Sequences of five potential recombination sites encoded close to an Immunoglobulin kappa constant region gene. *PNAS* **76**, 3450-3454.

McBlane, J. F., van Gent, D. C., Ramsden, D. A., Romeo, C., Cuomo, C. A., Gellert, M. and Oettinger, M. A. (1995). Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell* **83**, 387-395.

McCormack, W. T., Tjoelker, L. W., Stella, G., Postema, C. E. and Thompson, C. B. (1991). Chicken T-cell receptor beta-Chain diversity: An evolutionarily conserved Dbeta-encoded glycine turn within the hypervariable CDR3 Domain. *PNAS* **88**, 7699-7703.

McHeyzer-Williams, M. G. and Davis, M. M. (1995). Antigen-specific development of primary and memory T cells *in vivo*. *Science* **268**, 106-111.

McKeever, D. J., Taracha, E. L., Innes, E. L., MacHugh, N. D., Awino, E., Goddeeris, B. M. and Morrison, W. I. (1994). Adoptive transfer of immunity to *Theileria parva* in the CD8+ fraction of responding efferent lymph. *Proc Natl Acad Sci U S A* **91**, 1959-63.

McKeever, D. J., Nyanjui, J. K. and Ballingall, K. T. (1997). *In vitro* infection with *Theileria parva* is associated with IL10 expression in all bovine lymphocyte lineages. *Parasite Immunol* **19**, 319-24.

McKeever, D. J. and Morrison, W. I. (1998). Novel vaccines against *Theileria parva*: prospects for sustainability. *International Journal for Parasitology* **28**, 693-706.

McKeever, D. J., Taracha, E. L., Morrison, W. I., Musoke, A. J. and Morzaria, S. P. (1999). Protective immune mechanisms against *Theileria parva*: evolution of vaccine development strategies. *Parasitol Today* **15**, 263-7.

McKeever, D. J. (2006). *Theileria parva* and the bovine CTL response: down but not out? *Parasite Immunology* **28**, 339-345.

Mehlhorn, H. and Schein, E. (1984). The piroplasms: life cycle and sexual changes. *Advances in Parasitology* **23**, 37-103.

Meijers, R., Lai, C.-C., Yang, Y., Liu, J.-h., Zhong, W., Wang, J.-h. and Reinherz, E. L. (2005). Crystal structures of murine MHC class I H-2 Db and Kb molecules in complex with CTL epitopes from Influenza A virus: Implications for TCR repertoire selection and immunodominance. *Journal of Molecular Biology* **345**, 1099-1110.

Messaoudi, I., Patino, J. A. G., Dyall, R., LeMaoult, J., Nikolich, Zcaron and ugich, J. (2002). Direct link between MHC polymorphism, T-cell avidity, and diversity in immune defense. *Science* **298**, 1797-1800.

Meyer-Olson, D., Shoukry, N. H., Brady, K. W., Kim, H., Olson, D. P., Hartman, K., Shintani, A. K., Walker, C. M. and Kalams, S. A. (2004). Limited T-cell receptor diversity of HCV-specific T-cell responses is associated with CTL escape. *J. Exp. Med.* **200**, 307-319.

Meylan, F., De Smedt, M., Leclercq, G., Plum, J., Leupin, O., Marguerat, S. and Conrad, B. (2005). Negative thymocyte selection to HERV-K18 superantigens in humans. *Blood* **105**, 4377-4382.

Miccoli, M. C., Antonacci, R., Vaccarelli, G., Lanave, C., Massari, S., Cribiu, E. P. and Ciccarese, S. (2003). Evolution of TRG clusters in cattle and sheep genomes as drawn from the structural analysis of the ovine TRG2@ locus. *J Mol Evol* **57**, 52-62.

Miles, J. J., Elhassen, D., Borg, N. A., Silins, S. L., Tynan, F. E., Burrows, J. M., Purcell, A. W., Kjer-Nielsen, L., Rossjohn, J., Burrows, S. R. *et al.* (2005). CTL recognition of a bulged viral peptide involves biased TCR selection. *J Immunol* **175**, 3826-34.

Milicic, A., Price, D. A., Zimbwa, P., Booth, B. L., Brown, H. L., Easterbrook, P. J., Olsen, K., Robinson, N., Gileadi, U., Sewell, A. K. *et al.* (2005). CD8+ T-cell epitope-flanking mutations disrupt proteasomal processing of HIV-1 Nef. *J Immunol* **175**, 4618-26.

Minami, T., Spooner, P. R., Irvin, A. D., Ocama, J. G., Dobbelaere, D. A. and Fujinaga, T. (1983). Characterisation of stocks of *Theileria parva* by monoclonal antibody profiles. *Res Vet Sci* **35**, 334-40.

Minjauw, B., Rushton, J., James, A. D. and Upton, M. (1999). Financial analysis of East Coast Fever control strategies in traditionally managed Sanga cattle in Central Province of Zambia. *Preventive Veterinary Medicine* **38**, 35-45.

Mizushima, N., Kohsaka, H., Nanki, T., Ollier, W. E., Carson, D. A. and Miyasaka, N. (1997). HLA-dependent peripheral T-cell receptor (TCR) repertoire formation and its modification by rheumatoid arthritis (RA). *Clin Exp Immunol* **110**, 428-33.

Mo, A. X., van Lelyveld, S. F., Craiu, A. and Rock, K. L. (2000). Sequences that flank subdominant and cryptic epitopes influence the proteolytic generation of MHC class I-presented peptides. *J Immunol* **164**, 4003-10.

Mo, X. Y., Cascio, P., Lemerise, K., Goldberg, A. L. and Rock, K. (1999). Distinct proteolytic processes generate the C and N termini of MHC class I-binding peptides. *J Immunol* **163**, 5851-9.

Monaco, J. J., Cho, S. and Attaya, M. (1990). Transport protein genes in the murine MHC: possible implications for antigen processing. *Science* **250**, 1723-1726.

Montalbano, A., Ogwaro, K. M., Tang, A., Matthews, A. G. W., Larijani, M., Oettinger, M. A. and Feeney, A. J. (2003). V(D)J Recombination Frequencies Can Be Profoundly Affected by Changes in the Spacer Sequence. *J Immunol* **171**, 5296-5304.

Morel, S., Levy, F., Burlet-Schiltz, O., Brasseur, F., Probst-Kepper, M., Peitrequin, A. L., Monsarrat, B., Van Velthoven, R., Cerottini, J. C., Boon, T. *et al.* (2000). Processing of some antigens by the standard proteasome but not by the immunoproteasome results in poor presentation by dendritic cells. *Immunity* **12**, 107-17.

Morrison, W. I., Goddeeris, B. M., Teale, A. J., Groocock, C. M., Kemp, S. J. and Stagg, D. A. (1987). Cytotoxic T-cells elicited in cattle challenged with *Theileria parva* (Muguga): evidence for restriction by class I MHC determinants and parasite strain specificity. *Parasite Immunol* **9**, 563-78.

Morrison, W. I., Buscher, G., Murray, M., Emery, D. L., Masake, R. A., Cook, R. H. and Wells, P. W. (1981). *Theileria parva*: kinetics of infection in the lymphoid system of cattle. *Exp Parasitol* **52**, 248-60.

Morrison, W. I. and Davis, W. C. (1991). Differentiation antigens expressed predominantly on CD4- CD8- T lymphocytes (WC1, WC2). *Veterinary Immunology and Immunopathology* **27**, 71-76.

Morrison, W. I. (1996a). Influence of host and parasite genotypes on immunological control of Theileria parasites. *Parasitology* **112 Suppl**, S53-66.

Morrison, W. I. (1996b). Subunit vaccines for complex tick-borne pathogens: dream or reality? *Br Vet J* **152**, 617-9.

Morzaria, S. P., Irvin, A. D., Taracha, E., Spooner, P. R., Voigt, W. P., Fujinaga, T. and Katende, J. (1987). Immunization against East Coast fever: the use of selected stocks of *Theileria parva* for immunization of cattle exposed to field challenge. *Vet Parasitol* **23**, 23-41.

Morzaria, S. P., Irvin, A. D., Wathanga, J., D'Souza, D., Katende, J., Young, A. S., Scott, J. and Gettinby, G. (1988). The effect of East Coast fever immunisation and different acaricidal treatments on the productivity of beef cattle. *Vet Rec.* **123**, 313-320.

Morzaria, S. P., Spooner, P. R., Bishop, R. P., Musoke, A. J. and Young, J. R. (1990). SfiI and NotI polymorphisms in Theileria stocks detected by pulsed field gel electrophoresis. *Mol Biochem Parasitol* **40**, 203-11.

Moss, P. A. H., Moots, R. J., Rosenberg, W. M. C., Rowland-Jones, S. J., Bodmer, H. C., McMichael, A. J. and Bell, J. I. (1991). Extensive conservation of alpha and beta chains of the human T-cell antigen receptor recognizing HLA-A2 and Influenza A matrix peptide. *PNAS* **88**, 8987-8990.

Mostoslavsky, R., Alt, F. W. and Rajewsky, K. (2004). The lingering enigma of the allelic exclusion mechanism. *Cell* **118**, 539-544.

Moudgil, K. D., Sercarz, E. E. and Grewal, I. S. (1998). Modulation of the immunogenicity of antigenic determinants by their flanking residues. *Immunol Today* **19**, 217-20.

Muhammed, S. I., Lauerman, L. H., Jr. and Johnson, L. W. (1975). Effect of humoral antibodies on the course of *Theileria parva* infection (East Coast fever) of cattle. *Am J Vet Res* **36**, 399-402.

Musoke, A. J., Nantulya, V. M., Buscher, G., Masake, R. A. and Otim, B. (1982). Bovine immune response to *Theileria parva*: neutralizing antibodies to sporozoites. *Immunology* **45**, 663-8.

Musoke, A. J., Nantulya, V. M., Rurangirwa, F. R. and Buscher, G. (1984). Evidence for a common protective antigenic determinant on sporozoites of several *Theileria parva* strains. *Immunology* **52**, 231-8.

Mukhebi, A. W., Perry, B. D. and Kruska, R. (1992). Estimated economics of theileriosis control in Africa. *Preventive Veterinary Medicine* **12**, 73-85.

Musoke, A., Morzaria, S., Nkonge, C., Jones, E. and Nene, V. (1992). A recombinant sporozoite surface antigen of *Theileria parva* induces protection in cattle. *Proc Natl Acad Sci U S A* **89**, 514-8.

Musoke, A., Rowlands, J., Nene, V., Nyanjui, J., Katende, J., Spooner, P., Mwaura, S., Odongo, D., Nkonge, C., Mbogo, S. *et al.* (2005). Subunit vaccine based on the p67 major surface protein of *Theileria parva* sporozoites reduces severity of infection derived from field tick challenge. *Vaccine* **23**, 3084-95.

Mutugi, J. J., Young, A. S., Maritim, A. C., Ndungu, S. G., Mining, S. K., Linyonyi, A., Ngumi, P. N., Leitch, B. L., Morzaria, S. P. and Dolan, T. T. (1989). Immunisation of cattle against theileriosis in Coast Province, Kenya: laboratory evaluation of a *Theileria parva parva* stabilate for use in 'infection and treatment' immunisation in the field. *Res Vet Sci* **47**, 170-7.

Nadel, B., Tang, A., Escuro, G., Lugo, G. and Feeney, A. J. (1998a). Sequence of the spacer in the recombination signal sequence affects V(D)J rearrangement frequency and correlates with nonrandom Vkappa usage *in vivo*. *J Exp Med* **187**, 1495-503.

Nadel, B., Tang, A., Lugo, G., Love, V., Escuro, G. and Feeney, A. J. (1998b). Decreased frequency of rearrangement due to the synergistic effect of nucleotide changes in the heptamer and nonamer of the recombination signal sequence of the V kappa gene A2b, which is associated with increased susceptibility of Navajos to Haemophilus influenzae type b disease. *J Immunol* **161**, 6068-73.

Naessens, J., Newson, J., Williams, D. J. and Lutje, V. (1988). Identification of isotypes and allotypes of bovine immunoglobulin M with monoclonal antibodies. *Immunology* **63**, 569-74.

Nanda, N. K., Apple, R. and Sercarz, E. (1991). Limitations in plasticity of the T-cell receptor repertoire. *PNAS* **88**, 9503-9507.

Nanki, T., Kohsaka, H. and Miyasaka, N. (1998). Development of human peripheral TCRBJ gene repertoire. *J Immunol* **161**, 228-233.

Naumov, Y. N., Hogan, K. T., Naumova, E. N., Pagel, J. T. and Gorski, J. (1998). A class I MHC-restricted recall response to a viral peptide is highly polyclonal despite stringent CDR3 selection: implications for establishing memory T-cell repertoires in "real-world" conditions. *J Immunol* **160**, 2842-2852.

Neefjes, J. J., Momburg, F. and Hammerling, G. J. (1993). Selective and ATP-dependent translocation of peptides by the MHC-encoded transporter. *Science* **261**, 769-771.

Nelson, P. N., Hooley, P., Roden, D., Davari Ejtehadi, H., Rylance, P., Warren, P., Martin, J. and Murray, P. G. (2004). Human endogenous retroviruses: transposable elements with potential? *Clinical and Experimental Immunology* **138**, 1-9.

Nene, V., Iams, K. P., Gobright, E. and Musoke, A. J. (1992). Characterisation of the gene encoding a candidate vaccine antigen of *Theileria parva* sporozoites. *Mol Biochem Parasitol* **51**, 17-27.

Nene, V., Inumaru, S., McKeever, D., Morzaria, S., Shaw, M. and Musoke, A. (1995). Characterization of an insect cell-derived *Theileria parva* sporozoite vaccine antigen and immunogenicity in cattle. *Infect Immun* **63**, 503-8.

Nene, V., Musoke, A., Gobright, E. and Morzaria, S. (1996). Conservation of the sporozoite p67 vaccine antigen in cattle-derived *Theileria parva* stocks with different cross-immunity profiles. *Infect Immun* **64**, 2056-61.

Nikolich-Zugich, J., Slifka, M. K. and Messaoudi, I. (2004). The many important facets of T-cell repertoire diversity. *Nat Rev Immunol* **4**, 123-32.

Noonan, D. J., Kofler, R., Singer, P. A., Cardenas, G., Dixon, F. J. and Theofilopoulos, A. N. (1986). Delineation of a defect in T-cell receptor beta genes of NZW mice predisposed to autoimmunity. *J Exp Med* **163**, 644-53.

Norval, R. A., Perry, B. D. and Young, A. S. (1992). The Epidemiology of Theileriosis in Africa. London: Academic Press.

Nyangito, H. O., Richardson, J. W., Mukhebi, A. W., Mundy, D. S., Zimmel, P., Namken, J. and Perry, B. D. (1994). Whole farm economic analysis of East Coast fever immunization strategies in Kilifi District, Kenya. *Preventive Veterinary Medicine* **21**, 215-235.

Obata, F., Tsunoda, M., Ito, K., Ito, I., Kaneko, T., Pawelec, G. and Kashiwagi, N. (1993). A single universal primer for the T-cell receptor (TCR) variable genes enables enzymatic amplification and direct sequencing of TCRβ cDNA of various T-cell clones. *Human Immunology* **36**, 163-167.

Obexer-Ruff, G., Fluri, A., Hein, W., Lazary, S., Peterhans, E. and Bertoni, G. (1998). Caprine T-cell receptor variable beta-chain (TCRV beta) repertoire analysis and potential applications in cowdriosis immune response studies. *Ann N Y Acad Sci* **849**, 321-6.

Oestreich, K. J., Cobb, R. M., Pierce, S., Chen, J., Ferrier, P. and Oltz, E. M. (2006). Regulation of TCRβ gene assembly by a promoter/enhancer holocomplex. *Immunity* **24**, 381-391.

Olaru, A., Petrie, H. T. and Livak, F. (2005). Beyond the 12/23 rule of VDJ recombination independent of the Rag proteins. *J Immunol* **174**, 6220-6.

Ossendorp, F., Eggers, M., Neisig, A., Ruppert, T., Groettrup, M., Sijts, A., Mengede, E., Kloetzel, P. M., Neefjes, J., Koszinowski, U. *et al.* (1996). A single residue exchange within a viral CTL epitope alters proteasome-mediated degradation resulting in lack of antigen presentation. *Immunity* **5**, 115-24.

Oukka, M., Manuguerra, J. C., Livaditis, N., Tourdot, S., Riche, N., Vergnon, I., Cordopatis, P. and Kosmatopoulos, K. (1996). Protection against lethal viral infection by vaccination with nonimmunodominant peptides. *J Immunol* **157**, 3039-45.

Oura, C. A., Odongo, D. O., Lubega, G. W., Spooner, P. R., Tait, A. and Bishop, R. P. (2003). A panel of microsatellite and minisatellite markers for the characterisation of field isolates of *Theileria parva*. *Int J Parasitol* **33**, 1641-53.

Oura, C. A., Bishop, R., Wampande, E. M., Lubega, G. W. and Tait, A. (2004). The persistence of component *Theileria parva* stocks in cattle immunized with the 'Muguga cocktail' live vaccine against East Coast fever in Uganda. *Parasitology* **129**, 27-42.

Oura, C. A., Asiimwe, B. B., Weir, W., Lubega, G. W. and Tait, A. (2005). Population genetic analysis and sub-structuring of *Theileria parva* in Uganda. *Mol Biochem Parasitol* **140**, 229-39.

Padovan, E., Casorati, G., Dellabona, P., Meyer, S., Brockhaus, M. and Lanzavecchia, A. (1993). Expression of two T-cell receptor alpha chains: dual receptor T-cells. *Science* **262**, 422-424.

Padovan, E., Giachino, C., Cella, M., Valitutti, S., Acuto, O. and Lanzavecchia, A. (1995). Normal T lymphocytes can express two different T-cell receptor beta chains: implications for the mechanism of allelic exclusion. *J. Exp. Med.* **181**, 1587-1591.

Palmowski, M. J., Choi, E. M., Hermans, I. F., Gilbert, S. C., Chen, J. L., Gileadi, U., Salio, M., Van Pel, A., Man, S., Bonin, E. *et al.* (2002). Competition between CTL narrows the immune response induced by prime-boost vaccination protocols. *J Immunol* **168**, 4391-8.

Pannetier, C., Cochet, M., Darche, S., Casrouge, A., Zoller, M. and Kourilsky, P. (1993). The sizes of the CDR3 hypervariable regions of the murine T-cell receptor beta chains vary as a function of the Recombined germ-line segments. *PNAS* **90**, 4319-4323.

Pantaleo, G., Demarest, J. F., Soudeyns, H., Graziosi, C., Denis, F., Adelsberger, J. W., Borrow, P., Saag, M. S., Shaw, G. M., Sekalytt, R. P. *et al.* (1994). Major expansion of CD8+ T cells with a predominant Vbeta usage during the primary immune response to HIV. *Nature* **370**, 463-467.

Pantaleo, G., Demarest, James F., Schacker, T., Vaccarezza, M., Cohen, Oren J., Daucher, M., Graziosi, C., Schnittman, Steven S., Quinn, Thomas C., Shaw, George M. *et al.* (1997). The qualitative nature of the primary immune response to HIV infection is a prognosticator of disease progression independent of the initial level of plasma viremia. *PNAS* **94**, 254-258.

Pearson, T. W., Lundin, L. B., Dolan, T. T. and Stagg, D. A. (1979). Cell-mediated immunity to Theileria-transformed cell lines. *Nature* **281**, 678-80.

Peggs, K., Verfuerth, S., Pizzey, A., Ainsworth, J., Moss, P. and Mackinnon, S. (2002). Characterization of human cytomegalovirus peptide-specific CD8+ T-cell repertoire diversity following *in vitro* restimulation by antigen-pulsed dendritic cells. *Blood* **99**, 213-223.

Pegram, R. G., James, A. D., Bamhare, C., Dolan, T. T., Hove, T., Kanhai, G. K. and Latif, A. A. (1996). Effects of immunisation against *Theileria parva* on beef cattle productivity and economics of control options. *Trop Anim Health Prod* **28**, 99-111.

Pion, S., Christianson, G. J., Fontaine, P., Roopenian, D. C. and Perreault, C. (1999). Shaping the repertoire of cytotoxic T lymphocyte responses: Explanation for the immunodominance effect whereby cytotoxic T lymphocytes specific for immunodominant antigens prevent recognition of nondominant antigens. *Blood* **93**, 952-962.

Posnett, D. N., Vissinga, C. S., Pambuccian, C., Wei, S., Robinson, M. A., Kostyu, D. and Concannon, P. (1994). Level of human TCRBV3S1 (V beta 3) expression correlates with allelic polymorphism in the spacer region of the recombination signal sequence. *J Exp Med* **179**, 1707-11.

Price, D. A., West, S. M., Betts, M. R., Ruff, L. E., Brenchley, J. M., Ambrozak, D. R., Edghill-Smith, Y., Kuroda, M. J., Bogdan, D. and Kunstman, K. (2004). T-cell receptor recognition motifs govern immune escape patterns in acute SIV infection. *Immunity* **21**, 793-803.

Probst, H. C., Dumrese, T. and van den Broek, M. F. (2002). Cutting edge: competition for APC by CTLs of different specificities is not functionally important during induction of antiviral responses. *J Immunol* **168**, 5387-91.

Probst, J., Blumenthal, S. G., Tenzer, S., Weinschenk, T., Dittmer, J., Schoor, O., Six, A., Rammensee, H. G. and Pascolo, S. (2004). A conserved sequence in the mouse variable T cell receptor alpha recombination signal sequence 23-bp spacer can affect recombination. *Eur J Immunol* **34**, 2179-90.

Pullen, A. M., Marrack, P. and Kappler, J. W. (1988). The T-cell repertoire is heavily influenced by tolerance to polymorphic self-antigens. *Nature* **335**, 796-801.

Pullen, A. M., Potts, W., Wakeland, E. K., Kappler, J. and Marrack, P. (1990). Surprisingly uneven distribution of the T-cell receptor V beta repertoire in wild mice. *J Exp Med* **171**, 49-62.

Quinn, A., McInerney, M., Huffman, D., McInerney, B., Mayo, S., Haskins, K. and Sercarz, E. (2006). T-cells to a dominant epitope of GAD65 express a public CDR3 motif. *Int. Immunol.* **18**, 967-979.

Radley, D. E., Brown, C. G. D., Burridge, M. J., Cunningham, M. P., Kirimi, I. M., Purnell, R. E. and Young, A. S. (1975a). East coast fever: 1. Chemoprophylactic immunization of cattle against *Theileria parva* (Muguga) and five theilerial strains. *Veterinary Parasitology* **1**, 35-41.

Radley, D. E., Brown, C. G. D., Cunningham, M. P., Kimber, C. D., Musisi, F. L., Payne, R. C., Purnell, R. E., Stagg, S. M. and Young, A. S. (1975b). East coast fever: 3. Chemoprophylactic immunization of cattle using oxytetracycline and a combination of theilerial strains. *Veterinary Parasitology* **1**, 51-60.

Radley, D. E., Young, A. S., Brown, C. G. D., Burridge, M. J., Cunningham, M. P., Musisi, F. L. and Purnell, R. E. (1975c). East coast fever: 2. Cross-immunity trials with a Kenya strain of Theileria lawrencei. *Veterinary Parasitology* **1**, 43-50.

Ramsden, D. A., Baetz, K. and Wu, G. E. (1994). Conservation of sequence in recombination signal sequence spacers. *Nucl. Acids Res.* **22**, 1785-1796.

Richards, M. H. and Nelson, J. L. (2000). The Evolution of Vertebrate Antigen Receptors: A Phylogenetic Approach. *Mol Biol Evol* **17**, 146-155.

Robbins, P. F., Dudley, M. E., Wunderlich, J., El-Gamil, M., Li, Y. F., Zhou, J., Huang, J., Powell, D. J., Jr. and Rosenberg, S. A. (2004). Cutting Edge: Persistence of Transferred Lymphocyte Clonotypes Correlates with Cancer Regression in Patients Receiving Cell Transfer Therapy. *J Immunol* **173**, 7125-7130.

Robinson, J., Waller, M. J., Parham, P., Groot, N. d., Bontrop, R., Kennedy, L. J., Stoehr, P. and Marsh, S. G. E. (2003). IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. *Nucl. Acids Res.* **31**, 311-314.

Robinson, M. A., Mitchell, M. P., Wei, S., Day, C. E., Zhao, T. M. and Concannon, P. (1993). Organization of human T-cell receptor beta-chain genes: clusters of V beta genes are present on chromosomes 7 and 9. *Proc Natl Acad Sci U S A* **90**, 2433-7.

Rock, K. L., York, I. A., Saric, T. and Goldberg, A. L. (2002). Protein degradation and the generation of MHC class I-presented peptides. *Adv Immunol* **80**, 1-70.

Rock, K. L., York, I. A. and Goldberg, A. L. (2004). Post-proteasomal antigen processing for major histocompatibility complex class I presentation. *Nat Immunol* **5**, 670-677.

Rodriguez, F., Harkins, S., Slifka, M. K. and Whitton, J. L. (2002). Immunodominance in virus-induced CD8+ T-cell responses is dramatically modified by DNA immunization and is regulated by gamma interferon. *J Virol* **76**, 4251-9.

Rosenberg, W. M., Moss, P. A. and Bell, J. I. (1992). Variation in human T-cell receptor V beta and J beta repertoire: analysis using anchor polymerase chain reaction. *Eur J Immunol* **22**, 541-9.

Rowen, L., Koop, B. F. and Hood, L. (1996). The complete 685-Kilobase DNA sequence of the human beta T-cell receptor locus. *Science* **272**, 1755-1762.

Rudolph, M. G., Stanfield, R. L. and Wilson, I. A. (2006). How TCRS bind MHCs, peptides and co-recpetors. *Annual Review of Immunology* **24**, 419-466.

Rufer, N. (2005). Molecular tracking of antigen-specific T-cell clones during immune responses. *Current Opinion in Immunology* **17**, 441-447.

Ryu, C. J., Haines, B. B., Lee, H. R., Kang, Y. H., Draganov, D. D., Lee, M., Whitehurst, C. E., Hong, H. J. and Chen, J. (2004). The T-cell receptor beta variable gene promoter is required for efficient Vbeta rearrangement but not allelic exclusion. *Mol. Cell. Biol.* **24**, 7015-7023.

Sakano, H., Huppi, K., Heinrich, G. and Tonegawa, S. (1979). Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature* **280**, 288-294.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). In Molecular Cloning: A Laboratory Manual: Cold Spring Harbour Laboratory Press, USA.

Sandberg, J. K., Fast, N. M. and Nixon, D. F. (2001). Functional heterogeneity of cytokines and cytolytic effector molecules in human CD8+ T lymphocytes. *J Immunol* **167**, 181-187.

Sandberg, J. K., Grufman, P., Wolpert, E. Z., Franksson, L., Chambers, B. J. and Karre, K. (1998). Superdominance among immunodominant H-2Kb-restricted epitopes and reversal by dendritic cell-mediated antigen delivery. *J Immunol* **160**, 3163-9.

Sant'Angelo, D. B., Waterbury, P. G., Cohen, B. E., Martin, W. D., Van Kaer, L., Hayday, A. C. and Janeway, C. A., Jr. (1997). The imprint of intrathymic self-peptides on the mature T-cell receptor repertoire. *Immunity* **7**, 517-24.

Santra, S., Barouch, D. H., Kuroda, M. J., Schmitz, J. E., Krivulka, G. R., Beaudry, K., Lord, C. I., Lifton, M. A., Wyatt, L. S., Moss, B. *et al.* (2002). Prior vaccination increases the epitopic breadth of the cytotoxic T-lymphocyte response that evolves in rhesus monkeys following a simian-human immunodeficiency virus infection. *J Virol* **76**, 6376-81.

Saric, T., Chang, S. C., Hattori, A., York, I. A., Markant, S., Rock, K. L., Tsujimoto, M. and Goldberg, A. L. (2002). An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nat Immunol* **3**, 1169-76.

Savage, P. A., Boniface, J. J. and Davis, M. M. (1999). A kinetic basis for T-cell receptor repertoire selection during an immune response. *Immunity* **10**, 485-492.

Schatz, D. G. (2004). V(D)J recombination. *Immunological Reviews* **200**, 5-11.

Schein, E., Warnecke, M. and Kirmse, P. (1977). Development of *Theileria parva* (Theiler, 1904) in the gut of *Rhipicephalus appendiculatus* (Neumann, 1901). *Parasitology* **75**, 309-16.

Schlissel, M. S. (2003). Regulating antigen-receptor gene assembly. *Nature Reviews Immunology* **3**, 890-899.

Schrenzel, M. D., Watson, J. L. and Ferrick, D. A. (1994). Characterization of horse (*Equus caballus*) T-cell receptor beta chain genes. *Immunogenetics* **40**, 135-44.

Schultz, E. S., Chapiro, J., Lurquin, C., Claverol, S., Burlet-Schiltz, O., Warnier, G., Russo, V., Morel, S., Levy, F., Boon, T. *et al.* (2002). The production of a new MAGE-3 peptide presented to cytolytic T lymphocytes by HLA-B40 requires the immunoproteasome. *J Exp Med* **195**, 391-9.

Schwarz, K., van Den Broek, M., Kostka, S., Kraft, R., Soza, A., Schmidtke, G., Kloetzel, P. M. and Groettrup, M. (2000). Overexpression of the proteasome subunits LMP2, LMP7, and MECL-1, but not PA28 alpha/beta, enhances the presentation of an immunodominant lymphocytic choriomeningitis virus T cell epitope. *J Immunol* **165**, 768-78.

Seboun, E., Robinson, M. A., Kindt, T. J. and Hauser, S. L. (1989). Insertion/deletion-related polymorphisms in the human T-cell receptor beta gene complex. *J Exp Med* **170**, 1263-70.

Serwold, T., Gaw, S. and Shastri, N. (2001). ER aminopeptidases generate a unique pool of peptides for MHC class I molecules. *Nat Immunol* **2**, 644-51.

Serwold, T., Gonzalez, F., Kim, J., Jacob, R. and Shastri, N. (2002). ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature* **419**, 480-3.

Sette, A., Vitiello, A., Reherman, B., Fowler, P., Nayersina, R., Kast, W. M., Melief, C. J., Oseroff, C., Yuan, L., Ruppert, J. *et al.* (1994). The relationship between class I binding affinity and immunogenicity of potential cytotoxic T-cell epitopes. *J Immunol* **153**, 5586-92.

Sewell, A. K., Price, D. A., Teisserenc, H., Booth, B. L., Jr., Gileadi, U., Flavin, F. M., Trowsdale, J., Phillips, R. E. and Cerundolo, V. (1999). IFN-gamma exposes a cryptic cytotoxic T lymphocyte epitope in HIV-1 reverse transcriptase. *J Immunol* **162**, 7075-9.

Sha, W. C., Nelson, C. A., Newberry, R. D., Kranzt, D. M., Russell, J. H. and Loh, D. Y. (1988). Selective expression of an antigen receptor on CD8-bearing T lymphocytes in transgenic mice. *Nature* **335**, 271-274.

Shastri, N., Serwold, T. and Gonzalez, F. (1995). Presentation of endogenous peptide/MHC class I complexes is profoundly influenced by specific C-terminal flanking residues. *J Immunol* **155**, 4339-46.

Shatry, A. M., Wilson, A. J., Varma, S. and Dolan, T. T. (1981). Sequential study of lymph node and splenic aspirates during *Theileria parva* infection in calves. *Res Vet Sci* **30**, 181-4.

Shaw, M. K., Tilney, L. G. and Musoke, A. J. (1991). The entry of *Theileria parva* sporozoites into bovine lymphocytes: evidence for MHC class I involvement. *J Cell Biol* **113**, 87-101.

Shaw, M. K. and Young, A. S. (1995). Differential development and emission of *Theileria parva* sporozoites from the salivary gland of *Rhipicephalus appendiculatus. Parasitology* **111**, 153-60.

Shaw, M. K. (2003). Cell invasion by Theileria sporozoites. *Trends Parasitol* **19**, 2-6.

Shen, D.-F., Doukhan, L., Kalams, S. and Delwart, E. (1998). High-resolution analysis of T-cell receptor beta chain repertoires using DNA heteroduplex tracking: generally stable, clonal CD8+ expansions in all healthy young adults. *Journal of Immunological Methods* **215**, 113-121.

Sieh, P. and Chen, J. (2001). Distinct control of the frequency and allelic exclusion of the Vbeta gene rearrangement at the TCRbeta locus. *J Immunol* **167**, 2121-2129.

Sijts, A. J., Ruppert, T., Rehermann, B., Schmidt, M., Koszinowski, U. and Kloetzel, P. M. (2000). Efficient generation of a hepatitis B virus cytotoxic T lymphocyte epitope requires the structural features of immunoproteasomes. *J Exp Med* **191**, 503-14.

Sikes, M. L., Gomez, R. J., Song, J. and Oltz, E. M. (1998). A developmental stage-specific promoter directs germline transcription of Dbeta Jbeta gene segments in precursor T lymphocytes. *J Immunol* **161**, 1399-1405.

Siu, G., Clark, S. P., Yoshikai, Y., Malissen, M., Yanagi, Y., Strauss, E., Mak, T. W. and Hood, L. (1984). The human T-cell antigen receptor is encoded by variable, diversity, and joining gene segments that rearrange to generate a complete V gene. *Cell* **37**, 393-401.

Siu, G., Strauss, E. C., Lai, E. and Hood, L. E. (1986). Analysis of a human V beta gene subfamily. *J. Exp. Med.* **164**, 1600-1614.

Slifka, M. K., Blattman, J. N., Sourdive, D. J. D., Liu, F., Huffman, D. L., Wolfe, T., Hughes, A., Oldstone, M. B. A., Ahmed, R. and von Herrath, M. G. (2003). Preferential escape of subdominant CD8+ T-cells during negative selection results in an altered antiviral T-cell hierarchy. *J Immunol* **170**, 1231-1239.

Smith, L. R., Plaza, A., Singer, P. A. and Theofilopoulos, A. N. (1990). Coding sequence polymorphisms among V beta T-cell receptor genes. *J Immunol* **144**, 3234-3237.

Smith, L. R., Kono, D. H. and Theofilopoulos, A. N. (1991). Complexity and sequence identification of 24 rat V beta genes. *J Immunol* **147**, 375-9.

Sourdive, D. J. D., Murali-Krishna, K., Altman, J. D., Zajac, A. J., Whitmire, J. K., Pannetier, C., Kourilsky, P., Evavold, B., Sette, A. and Ahmed, R. (1998). Conserved T-cell receptor repertoire in primary and memory CD8+ T-cell responses to an acute viral infection. *J. Exp. Med.* **188**, 71-82.

Spies, T. and DeMars, R. (1991). Restored expression of major histocompatibility class I molecules by gene transfer of a putative peptide transporter. *Nature* **351**, 323-324.

Stagg, D. A., Dolan, T. T., Leitch, B. L. and Young, A. S. (1981). The initial stages of infection of cattle cells with *Theileria parva* sporozoites *in vitro*. *Parasitology* **83**, 191-7.

Starr, T. K., Jameson, S. C. and Hogquist, K. A. (2003). Positive and negative selection of T-cells. *Annual Review of Immunology* **21**, 139-176.

Stewart-Jones, G. B. E., McMichael, A. J., Bell, J. I., Stuart, D. I. and Jones, E. Y. (2003). A structural basis for immunodominant human T-cell receptor recognition. *Nat Immunol* **4**, 657-663.

Stoltze, L., Schirle, M., Schwarz, G., Schroter, C., Thompson, M. W., Hersh, L. B., Kalbacher, H., Stevanovic, S., Rammensee, H. G. and Schild, H. (2000). Two new proteases in the MHC class I processing pathway. *Nat Immunol* **1**, 413-8.

Storset, A. K., S, K., Berg, I., Boysen, P., Hope, J. C. and Dissen, E. (2004). NKp46 defines a subset of bovine leukocytes with natural killer cell characteristics. *European Journal of Immunology* **34**, 669-676.

Su, C. and Nei, M. (2001). Evolutionary dynamics of the T-cell receptor VB gene family as inferred from the human and mouse genomic sequences. *Mol Biol Evol* **18**, 503-513.

Tan, L. C., Gudgeon, N., Annels, N. E., Hansasuta, P., O'Callaghan, C. A., Rowland-Jones, S., McMichael, A. J., Rickinson, A. B. and Callan, M. F. C. (1999). A re-evaluation of the frequency of CD8+ T-cells specific for EBV in healthy virus carriers. *J Immunol* **162**, 1827-1835.

Tanaka, A., Ishiguro, N. and Shinagawa, M. (1990). Sequence and diversity of bovine T-cell receptor beta-chain genes. *Immunogenetics* **32**, 263-71.

Taracha, E. L., Goddeeris, B. M., Scott, J. R. and Morrison, W. I. (1992). Standardization of a technique for analysing the frequency of parasite-specific cytotoxic T lymphocyte precursors in cattle immunized with *Theileria parva*. *Parasite Immunol* ***14****, 143-54.*

Taracha, E. L., Goddeeris, B. M., Morzaria, S. P. and Morrison, W. I. (1995a). Parasite strain specificity of precursor cytotoxic T-cells in individual animals correlates with cross-protection in cattle challenged with *Theileria parva*. *Infect Immun* **63**, 1258-62.

Taracha, E. L., Goddeeris, B. M., Teale, A. J., Kemp, S. J. and Morrison, W. I. (1995b). Parasite strain specificity of bovine cytotoxic T-cell responses to *Theileria parva* is determined primarily by immunodominance. *J Immunol* **155**, 4854-60.

Taracha, E. L., Awino, E. and McKeever, D. J. (1997). Distinct CD4+ T-cell helper requirements in *Theileria parva*-immune and -naive bovine CTL precursors. *J Immunol* **159**, 4539-45.

Taylor, A. H., Haberman, A. M., Gerhard, W. and Caton, A. J. (1990). Structure-function relationships among highly diverse T-cells that recognize a determinant from influenza virus hemagglutinin. *J. Exp. Med.* **172**, 1643-1651.

Teng, M. K., Smolyar, A., Tse, A. G. D., Liu, J. H., Liu, J., Hussey, R. E., Nathenson, S. G., Chang, H. C., Reinherz, E. L. and Wang, J. H. (1998). Identification of a common docking topology with substantial variation among different TCR-peptide-MHC complexes. *Current Biology* **8**, 409-414.

Tenzer, S., Stoltze, L., Schonfisch, B., Dengjel, J., Muller, M., Stevanovic, S., Rammensee, H. G. and Schild, H. (2004). Quantitative analysis of prion-protein degradation by constitutive and immuno-20S proteasomes indicates differences correlated with disease susceptibility. *J Immunol* **172**, 1083-91.

Thatcher, T. H., O'Brien, D. P., Altuwaijri, S. and Barth, R. K. (2000). Increasing the frequency of T-cell precursors specific for a cryptic epitope of hen-egg lysozyme converts it to an immunodominant epitope. *Immunology* **99**, 235-42.

Toes, R. E., Nussbaum, A. K., Degermann, S., Schirle, M., Emmerich, N. P., Kraft, M., Laplace, C., Zwinderman, A., Dick, T. P., Muller, J. ***et al.*** (2001). Discrete cleavage motifs of constitutive and immunoproteasomes revealed by quantitative analysis of cleavage products. *J Exp Med* **194**, 1-12.

Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature* **302**, 575-581.

Trautmann, L., Rimbert, M., Echasserieau, K., Saulquin, X., Neveu, B., Dechanet, J., Cerundolo, V. and Bonneville, M. (2005). Selection of T-cell clones expressing high-affinity public TCRs within human cytomegalovirus-specific CD8+ T-cell responses. *J Immunol* **175**, 6123-6132.

Turnbull, E. L., Lopes, A. R., Jones, N. A., Cornforth, D., Newton, P., Aldam, D., Pellegrino, P., Turner, J., Williams, I., Wilson, C. M. ***et al.*** (2006). HIV-1 epitope-specific CD8+ T-cell responses strongly associated with delayed disease progression cross-recognize epitope variants efficiently. *J Immunol* **176**, 6130-6146.

Turner, S. J., Diaz, G., Cross, R. and Doherty, P. C. (2003). Analysis of clonotype distribution and persistence for an Influenza virus-specific CD8+ T-cell response. *Immunity* **18**, 549-559.

Turner, S. J., Kedzierska, K., Komodromou, H., La Gruta, N. L., Dunstone, M. A., Webb, A. I., Webby, R., Walden, H., Xie, W., McCluskey, J. *et al.* (2005). Lack of prominent peptide-major histocompatibility complex features limits repertoire diversity in virus-specific CD8+ T-cell populations. *Nat Immunol* **6**, 382-389.

Uematsu, Y. (1991). A novel and rapid cloning method for the T-cell receptor variable region sequences. *Immunogenetics* **34**, 174-8.

Utz, U., Banks, D., Jacobson, S. and Biddison, W. E. (1996). Analysis of the T-cell receptor repertoire of human T-cell leukemia virus type 1 (HTLV-1) Tax-specific CD8+ cytotoxic T lymphocytes from patients with HTLV-1-associated disease: evidence for oligoclonal expansion. *J. Virol.* **70**, 843-851.

Valmori, D., Gervois, N., Rimoldi, D., Fonteneau, J.-F., Bonelo, A., Lienard, D., Rivoltini, L., Jotereau, F., Cerottini, J.-C. and Romero, P. (1998). Diversity of the fine specificity displayed by HLA-A*0201-restricted CTL specific for the immunodominant Melan-A/MART-1 antigenic peptide. *J Immunol* **161**, 6956-6962.

van der Burg, S. H., Visseren, M. J., Brandt, R. M., Kast, W. M. and Melief, C. J. (1996). Immunogenicity of peptides bound to MHC class I molecules depends on the MHC-peptide complex stability. *J Immunol* **156**, 3308-14.

van der Most, R. G., Sette, A., Oseroff, C., Alexander, J., Murali-Krishna, K., Lau, L. L., Southwood, S., Sidney, J., Chesnut, R. W., Matloubian, M. *et al.* (1996). Analysis of cytotoxic T-cell responses to dominant and subdominant epitopes during acute and chronic lymphocytic choriomeningitis virus infection. *J Immunol* **157**, 5543-54.

Vijh, S., Pilip, I. M. and Pamer, E. G. (1998). Effect of antigen-processing efficiency on *in vivo* T cell response magnitudes. *J Immunol* **160**, 3971-7.

Vissinga, C. S., Charmley, P. and Concannon, P. (1994). Influence of coding region polymorphism on the peripheral expression of a human TCR V beta gene. *J Immunol* **152**, 1222-7.

von Boehmer, H., Aifantis, I., Gounari, F., Azogui, O., Haughn, L., Apostolou, I., Jaeckel, E., Grassi, F. and Klein, L. (2003). Thymic selection revisited: how essential is it? *Immunological Reviews* **191**, 62-78.

Wade, T., Bill, J., Marrack, P. C., Palmer, E. and Kappler, J. W. (1988). Molecular basis for the nonexpression of V beta 17 in some strains of mice. *J Immunol* **141**, 2165-7.

Wagner, G. G., Duffus, W. P. and Burridge, M. J. (1974). The specific immunoglobulin response in cattle immunized with isolated *Theileria parva* antigens. *Parasitology* **69**, 43-53.

Walker, B. D. and Korber, B. T. (2001). Immune control of HIV: the obstacles of HLA and viral diversity. *Nat Immunol* **2**, 473-475.

Wallace, M. E., Bryden, M., Cose, S. C., Coles, R. M., Schumacher, T. N., Brooks, A. and Carbone, F. R. (2000). Junctional biases in the naive TCR repertoire control the CTL response to an immunodominant determinant of HSV-1. *Immunity* **12**, 547-56.

Weaver, D. T. (1995). V(D)J recombination and double-strand break repair. *Adv Immunol* **58**, 29-85.

Webster, P., Dobbelaere, D. A. and Fawcett, D. W. (1985). The entry of sporozoites of *Theileria parva* into bovine lymphocytes *in vitro*. Immunoelectron microscopic observations. *Eur J Cell Biol* **36**, 157-62.

Wilde, J. K. (1966). Changes in bovine bone marrow during the course of East Coast fever. *Res Vet Sci* **7**, 213-24.

Williams, C. B., Blankenhorn, E. P., Byrd, K. E., Levinson, G. and Gutman, G. A. (1991). Organization and nucleotide sequence of the rat T-cell receptor beta-chain complex. *J Immunol* **146**, 4406-13.

Willis, R. A., Kappler, J. W. and Marrack, P. C. (2006). CD8+ T-cell competition for dendritic cells *in vivo* is an early event in activation. *PNAS* **103**, 12063-12068.

Wilson, A., Marechal, C. and MacDonald, H. R. (2001). Biased Vbeta usage in immature thymocytes is independent of DJbeta proximity and pTalpha pairing. *J Immunol* **166**, 51-57.

Wolpert, E. Z., Grufman, P., Sandberg, J. K., Tegnesjo, A. and Karre, K. (1998). Immunodominance in the CTL response against minor histocompatibility antigens: interference between responding T-cells, rather than with presentation of epitopes. *J Immunol* **161**, 4499-505.

Wotton, D., Ways, D. K., Parker, P. J. and Owen, M. J. (1993). Activity of both Raf and Ras is necessary for activation of transcription of the human T cell receptor beta gene by protein kinase C, Ras plays multiple roles. *J. Biol. Chem.* **268**, 17975-17982.

Wu, C., Bassing, C. H., Jung, D., Woodman, B. B., Foy, D. and Alt, F. W. (2003). Dramatically increased rearrangement and peripheral representation of Vbeta14 driven by the 3'Dbeta1 recombination signal sequence. *Immunity* **18**, 75-85.

Yewdell, J. W. and Bennink, J. R. (1999). Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses. *Annu Rev Immunol* **17**, 51-88.

Yewdell, J. W. and Del Val, M. (2004). Immunodominance in CD8+ responses to viruses: cell biology, cellular immunology, and mathematical models. *Immunity* **21**, 149-53.

York, I. A., Chang, S. C., Saric, T., Keys, J. A., Favreau, J. M., Goldberg, A. L. and Rock, K. L. (2002). The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8-9 residues. *Nat Immunol* **3**, 1177-84.

York, I. A., Brehm, M. A., Zendzian, S., Towne, C. F. and Rock, K. L. (2006). Endoplasmic reticulum aminopeptidase 1 (ERAP1) trims MHC class I-presented peptides *in vivo* and plays an important role in immunodominance. *PNAS* **103**, 9202-9207.

Yoshimura, Y., Yadav, R., Christianson, G. J., Ajayi, W. U., Roopenian, D. C. and Joyce, S. (2004). Duration of alloantigen presentation and avidity of T-cell antigen recognition correlate with immunodominance of CTL response to minor histocompatibility antigens. *J Immunol* **172**, 6666-74.

Young, A. S. and Leitch, B. L. (1980). A probable relationship between the development of Theileria species and the ecdysis of their tick hosts. *J Parasitol* **66**, 356-9.

Young, A. S. and Leitch, B. L. (1981). Epidemiology of East Coast fever: some effects of temperature on the development of *Theileria parva* in the tick vector, *Rhipicephalus appendiculatus*. *Parasitology* **83**, 199-211.

Young, A. S., Leitch, B. L., Stagg, D. A. and Dolan, T. (1983). Identification of Theileria infections in living salivary glands of ticks. *Parasitology* **86**, 519-528.

Young, A. S., Mutugi, J. J., Kariuki, D. P., Lampard, D., Maritim, A. C., Ngumi, P. N., Linyonyi, A., Leitch, B. L., Ndungu, S. G. and Lesan, A. C. (1992). Immunisation of cattle against theileriosis in Nakuru District of Kenya by infection and treatment and the introduction of unconventional tick control. *Veterinary Parasitology* **42**, 225-240.

Zhang, Q., Ragnauth, C., Greener, M. J., Shanahan, C. M. and Roberts, R. G. (2002). The Nesprins are giant actin-binding proteins, orthologous to *Drosophila melanogaster* muscle protein MSP-300. *Genomics* **80**, 473-481.

Zhao, T. M., Whitaker, S. E. and Robinson, M. A. (1994). A genetically determined insertion/deletion related polymorphism in human T-cell receptor beta chain (TCRB) includes functional variable gene segments. *J Exp Med* **180**, 1405-14.

Zhong, W. and Reinherz, E. L. (2004). *In vivo* selection of a TCR Vbeta repertoire directed against an immunodominant influenza virus CTL epitope. *Int. Immunol.* **16**, 1549-1559.

Zhou, J., Dudley, M. E., Rosenberg, S. A. and Robbins, P. F. (2004). Selective growth, in vitro and *in vivo*, of individual T-cell clones from tumor-infiltrating lymphocytes obtained from patients with melanoma. *J Immunol* **173**, 7622-7629.

Zhu, H., Liu, K., Cerny, J., Imoto, T. and Moudgil, K. D. (2005). Insertion of the dibasic motif in the flanking region of a cryptic self-determinant leads to activation of the epitope-specific T cells. *J Immunol* **175**, 2252-60.

Immunogenetics (2005) 57: 674–681
DOI 10.1007/s00251-005-0040-y

ORIGINAL PAPER

E. F. Houston · T. Connelley · K. Parsons ·
N. D. MacHugh · W. I. Morrison

Analysis of T-cell receptor *BV* gene sequences in cattle reveals extensive duplication within the *BV9* and *BV20* subgroups

Received: 30 March 2005 / Accepted: 8 August 2005 / Published online: 27 September 2005

Abstract We investigated the repertoire of functional T-cell receptor β-chain variable genes (*TRBV* genes) in cattle by analysing the nucleotide sequences and predicted amino acid sequences of a set of cDNA clones isolated from lymph node T cells. Thirty-nine distinct *TRBV* sequences were identified, bringing the total number of recognised bovine *TRBV* gene segments to more than 40. Sixteen *TRBV* sub groups were defined based on their sequence homology to each other and to human *TRBV* genes. All of the main phylogenetic lineages of *BV* gene subgroups described in humans and mice were represented. Eight of the subgroups were found to contain more than one member. The most striking feature of the results was the large number of sequences (more than half of the sequenced clones) in the *BV9* and *BV20* subgroups, which were found to contain 12 and 8 distinct sequences, respectively. In contrast, the corresponding human *TRBV* subfamilies contain a single member. The results indicate that, as in humans, there has been extensive gene duplication within the *TRBV* locus during evolution. However, duplication of different *BV* subgroups in cattle has resulted in a *TRBV* gene repertoire distinct from that found in other species.

Introduction

The generation of a repertoire of T lymphocytes with unique receptors for antigenic recognition is a key feature of the immune system of higher vertebrates which enables them to mount highly specific T-cell-mediated immune responses to a wide variety of pathogenic organisms. The T-cell receptor (TR) consists of a disulphide-linked heterodimer that specifically recognises peptides associated with major histocompatibility complex (MHC) proteins on the surface of antigen-presenting cells. The diversity displayed by the α and β (or γ and δ) chains that make up the heterodimer is determined by multiple variable (*V*), diversity (*D*; β- and δ-chains) and joining (*J*) genes that associate more or less randomly during somatic recombination to generate a variable exon immediately upstream of a constant (*C*) gene segment (Davis 1990).

The *TRBV* genes that encode the variable region of the TR β-chain are classified into subgroups, the members of which share >75% sequence homology at the nucleotide level (Arden et al. 1995a,b). Comparison of the sequences of *TRV* subgroups from different species reveals that most sequences have a much higher degree of similarity to a homologue in another species than to other subgroups in the same species. For example, when the deduced protein sequences of human and mouse *TRBV* genes were compared, 14 murine subgroups, which display ~40% identity, were found to have >60% identity to their human homologues (Clark et al. 1995). In addition, physical maps of the *TRB* locus in mice and humans have demonstrated that the relative order of the *TRBV* homologues is conserved (Lai et al. 1988; Rowen et al. 1996). However, both the number of *BV* subgroups and the number of functional gene segments within the *BV* subgroups have been shown to differ between humans and mice. While most murine subgroups comprise a single functional gene segment, some human genes have undergone extensive duplication, resulting in subgroups with multiple members. These observations indicate that a set of ancestral *V* gene segments has been retained through speciation and that different duplication and deletion events have resulted in the development of unique *TRBV* repertoires in different species.

Information on the complement of *TR* variable genes in humans and mice has allowed the development of methods to determine *TRBV* gene usage of responding T cells and to identify abundant T-cell clones based on unique features of

E. F. Houston · K. Parsons
Institute for Animal Health, Compton,
Newbury, Berkshire, RG20 7NN, UK

T. Connelley · N. D. MacHugh · W. I. Morrison (✉)
Division of Animal Health and Welfare,
Royal (Dick) School of Veterinary Studies,
Easter Bush Veterinary Centre, Roslin,
Midlothian, EH25 9RG, UK
e-mail: ivan.morrison@ed.ac.uk

the third complementarity determining region (CDR3) (Akatsuka et al. 1999; Bernardin et al. 2003). Such studies have provided useful information on the clonal composition of T-cell responses and long-term fate of T-cell clones (Cohen et al. 2002; Turner et al. 2003; Bernardin et al. 2004). However, these methodologies are not currently applicable in farm animal species because of limited information on their *TR* genes. In the case of cattle, small numbers of cDNA clones encoding the bovine TR α-, β-, γ- and δ-chains have been isolated and sequenced (Ishiguro et al. 1990; Takeuchi et al. 1992; Tanaka et al. 1990). A series of *TRB* sequences isolated from a lymphocyte cDNA library contained 13 distinct *TRBV* gene segments representing nine different subgroups (Tanaka et al. 1990). Based on these findings and the results of Southern blotting of genomic DNA, these workers suggested that the identified subgroups contain relatively few members.

Herein, we report the results of a more extensive analysis of bovine *TRBV* sequences, which identifies 39 distinct *TRBV* gene sequences, 20 of which fall within two multimember subgroups, *BV9* and *BV20*. These findings indicate that there have been multiple gene duplication events within the bovine *TRB* locus, leading to marked expansion of the *BV9* and *BV20* subgroups.

Materials and methods

Sources of RNA

RNA samples for analysis of *TRBV* gene sequences were obtained from lymphocytes isolated from the peripheral lymph nodes of two Holstein–Friesian cattle (3327 and 5158). Three populations of cells were used: (i) unfractionated mononuclear cells isolated from a prescapular lymph node removed surgically from animal 3327, (ii) activated ($CD3^+$, class II MHC^+) T cells isolated from the contralateral lymph node of the same animal 9 days after infection with the protozoan parasite *Theileria parva* and (iii) a $CD2^-CD8^+$ subset of T cells isolated from a prescapular lymph node of animal 5158, also 9 days after infection with *T. parva*. Infection with *T. parva* has been shown to result in the appearance of large numbers of non-parasitised $CD2^-$, $CD8^+$, $CD3^+$, $\gamma/\delta TR^-$ lymphoblasts in the lymph node draining the site of inoculation (Morrison and McKeever 1998). Purification of the T-cell subsets was achieved by two colour immunofluorescence staining followed by separation with a FACStar Plus cell sorter (Becton Dickinson).

Amplification of *TRBV* transcripts

Total RNA was extracted from lymph node cells using an Ultraspec RNA isolation kit (Biotecx, AMS Biotechnology UK Ltd., Witney, Oxon). cDNA containing *TRBV* sequences was derived by reverse transcription of RNA using Superscript RT II (Gibco BRL, Paisley, UK), with an anti-sense primer, PBC1 (5′-GAGATCTCTGCTTCCGAGG-3′), complementary to a conserved region near the 5′ end of the bovine *TRBC*. First-strand cDNA was purified using a GLASSMax spin cartridge kit (Gibco), and a homopolymer tail of cytosine bases was added to the 3′ end using terminal deoxynucleotidyl transferase. Tailed cDNA was amplified by anchored polymerase chain reaction (PCR) with primers complementary to the homopolymer tail (5′-ACTACTACA AGCTTGGGIIGGGIIGGGIIG-3′) and the 5′ end of bovine *TRBC* genes (PBC2; 5′-GTTCGAACACAGCCACCTTG-3′). Amplification conditions were as follows: denaturing at 94°C for 1 min, annealing at 55°C for 1 min and extension at 72°C for 1 min, for 35 cycles with a final extension step of 72°C for 10 min. PCR products of approximately 550 bp were excised from agarose gels. DNA recovered using a Geneclean II kit (BIO 101, Vista, CA, USA) was cloned into pCRII plasmid vector using the TA cloning kit (Invitrogen, R&D Systems, Abingdon, Oxon, UK) and transformed into OneShot INVαF′ competent *E. coli* (Invitrogen), allowing blue/white selection of colonies containing inserts.

Sequencing of *TRBV* inserts

Sequencing reactions were carried out by the chain termination method (Sanger et al. 1977) using PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems Inc., Warrington, Cheshire, UK) and primers complementary to vector sequences flanking the insert (5′-AGCTCGGATCCACTAGTAAC-3′ and 5′-TCTAGATGC ATGCTCGAGCG-3′). Electrophoresis of the extension products was performed on a 373A DNA Sequencer (Applied Biosystems). All clones containing rearranged *TRB* transcripts were sequenced at least twice.

Sequence analysis

Sequence data were analysed using the GAP, BESTFIT, FASTA, PILEUP and LINEUP programs from the University of Wisconsin Genetic Computing Group (GCG) package (Devereux et al. 1984). The sequences were edited to remove leader, *BD*, *BJ* and *BC* sequences. The 3′ ends of the *V* gene segments were assigned as three codons after the conserved cysteine residue at position 92, according to convention. The sequences were classified into *TRBV* subgroups by screening them for nucleotide sequence homology to human and available bovine *TRBV* sequences in the GenEMBL database, using the FASTA program. Sequences sharing greater than 75% identity at the nucleotide level are defined as being members of the same subgroup. The nomenclature used for classification of the *TRBV* sequences is based on the IMGT system established for human and mouse *BV* subgroups (Lefranc et al. 1998; Giudicelli and Lefranc 1999). Numbers assigned on the basis of previous nomenclature (Arden et al. 1995a) are also listed in parentheses in Table 1.

Table 1 Classification of bovine *TRBV* sequences into subgroups

TRBV subgroup[a]	Number of clones sequenced	Number of subgroup members	Subgroup members identified previously[b]	
			Identical	Different
BV3 (*BV9*)	1	1		
BV4 (*BV7*)	2	1	1	
BV6 (*BV13*)	7	3		
BV7 (*BV6*)	5	1		2
BV9 (*BV1*)	38	12	1	
BV14 (*BV16*)	1	1		
BV15 (*BV24*)	5	1		
BV19 (*BV17*)	2	1		1
BV20 (*BV2*)	23	8	2	
BV21 (*BV10*)	8	3		
BV24 (*BV15*)	3	2	1	
BV26 (*BV28*)[c]	2	1	1	
BV27 (*BV14*)	3	1		
BV28 (*BV3*)	2	1	1	
BV29 (*BV4*)	1	1		1
BV30 (*BV20*)	4	1		
Total	107	39	7	4

[a]Classification is based on the IMGT nomenclature described in Giudicelli and Lefranc (1999). Numbers assigned on the basis of previous nomenclature (Arden et al. 1995a) are given in parentheses

[b]All described in Tanaka et al. (1990) except one—*BV7* (*BV6* in previous nomenclature) sequence reported by Buitkamp et al. (1993)

[c]Previously classified by Tanaka et al. (1990) as *BV10* (*BV12* in previous nomenclature)

Southern blotting

Genomic DNA was extracted from primary cultures of bovine skin fibroblasts using Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA), and aliquots were incubated overnight at 37°C with the following enzymes: *Hin*dIII, *Xba*I (10 U/μg DNA), *Sph*I and *Ssp*I (2.5 U/μg DNA). The samples were run overnight on a 0.8% agarose slab gel, and the DNA fragments were transferred to a nitrocellulose filter (Hybond-N+, Amersham Biosciences, Chalfont St. Giles, UK), in 20× SSC buffer as described in Maniatis (1992). *BV9*, *BV20* and *BV27* gene segments amplified from cloned plasmids by PCR were purified with Wizard PCR Preps DNA Purification System (Promega, Madison, WI, USA) and labelled with ^{32}P-dCTP using a Random Primed DNA labelling kit (Boehringer Mannheim, Lewes, East Sussex, UK). The blots were hybridised with the radiolabelled probes in Rapid-Hyb (Amersham Biosciences, Little Chalfont, Bucks, UK), with an incubation period of 2 h at 65°C and final stringency washes in 0.1× SSC, 0.1% (w/v) SDS at 65°C. The blots were exposed to x ray films for 16 h at −70°C.

Results

Sequence analysis of *TRBV* cDNA clones

Sequence information was obtained from a total of 107 clones that contained rearranged *V–D–J–C* transcripts with a correct open reading frame. Forty-two of these clones were derived from the resting lymph node cells, 43 from the activated population from the same animal and 22 from the activated cells from the second animal. The 107 clones included four sequences that reoccurred (i.e. they were identical throughout the *V D J* region), including three copies of one sequence and three identical pairs. The duplicated sequences were found only in the clones derived from the activated T-cell populations, which would be more likely to include clonally expanded T cells and hence yield more than one copy of the same *V–D–J* rearrangement. Nevertheless, there was no obvious preferential usage of particular *BV* gene segments by either of the activated T cell populations ($CD3^+$, class II^+ or $CD2^-CD8^+$) from which cDNA was isolated (data not shown).

Identification of 16 *TRBV* subfamilies

The 107 cDNA clones that were sequenced were classified into 16 subgroups based on their nucleotide sequence homology with human *TRBV* sequences (Table 1). Five of these subgroups were found to contain more than one member, resulting in a total of 39 distinct *BV* gene sequences. The clones included sequences identical to the members of six subgroups reported by Tanaka et al. (1990) (clones referred to in the text below with the prefix BTB), namely *BV4*, *BV9*, *BV10*, *BV20*, *BV24* and *BV28*. However, the sequence reported as *BV10* (*BV12* in old nomenclature) by Tanaka et al. (1990) has been reclassified as *BV26* since it displayed 80% similarity with human *BV26* as compared with 66% similarity with human *BV10*. New individual subgroup members were identified for a further five previously reported subgroups (*BV7*, *BV9*, *BV19*, *BV20* and *BV29*) (Buitkamp et al. 1993; Tanaka et al. 1990). Seven new *TRBV* subgroups (*BV3*, *BV6*, *BV14*, *BV15*, *BV21*, *BV27* and *BV30*), whose members showed closer homology to human *BV* sequences than to previously identified bovine sequences, were defined; detailed information on the sequences of these genes is reported elsewhere (Houston and Morrison 1999).

The *BV9* and *BV20* subgroups contain multiple members

Sixty-one (57%) of the cloned *TRBV* sequences were found to fall within two multi-member subgroups, *BV9* and *BV20*.

The BV9 subgroup

Thirty eight clones (35%) were classified as belonging to the bovine *BV9* subgroup (Table 2). Twelve distinct sequences, comprising ten groups of two or more clones that were identical and two single clones, were identified (designated *BV9-1* to *BV9-12*). Two of the sequences, *BV9-3* and *BV9-4*, differed from each other by only four nucleotides, whereas the remaining sequences all differed by at least 12 nucleotides. Although it is possible that *BV9-3* and *BV9-4* represent alleles, in the absence of genetic data, they have tentatively been classified, along with the other sequences, as subgroup members. The four clones classified as *BV9-1* were identical to a *BV9* clone (BTB18) identified previously by Tanaka et al. (1990). However, the alignment of BTB18 with the nucleotide sequences of *BV9* subgroup members results in a frameshift in the BTB18 sequence between residues 165 and 191, which leads to a complete change in nine amino acids of the deduced protein sequence in this region and elimination of the residues Arg-64 and Phe-65, which are found in nearly all other members of the *BV9* subgroup and in homologous human sequences. This suggests that the anomaly is due to an error in the published sequence of BTB18. All 12 *BV9* sequences were represented in the cDNA clones derived from one of the animals in the study.

The BV20 subgroup

Twenty-three sequences were classified as belonging to the bovine *BV20* subgroup (Table 3), including sequences with close similarity to the *BV20* sequences (BTB1, BTB4 and BTB10) identified by Tanaka et al. (1990). Eight distinct sequences, comprising seven groups of two or more clones that were identical and an additional single clone, were identified (designated *BV20-1* to *BV20-8*). Two of the sequences (*BV20-3* and *BV20-4*) differed by only five nucleotides, whereas all the remaining sequences differed by at least nine nucleotides. Although *BV20-3* and *BV20-4* could be alleles, in the absence of genetic data, they have tentatively been classified, along with the remaining sequences, as subgroup members. Maximal alignment of the nucleotide sequences with the previously published *BV20* sequences requires a dinucleotide insertion near the 5′ end of the *V* gene in the published BTB1 and BTB4 sequences, resulting in a change in the predicted protein sequences. The predicted amino acid sequences for BTB1 and BTB4 do not contain valine or isoleucine at position 6 or proline at position 8, residues that are found in the deduced protein sequences of every *BV9* subgroup member identified in the present study and are conserved in almost every *TRBV* gene segment regardless of species. This suggests that the deletion of two nucleotides near the 5′ ends of BTB1 and BTB4 has arisen from defective mRNA transcripts or from errors in the sequencing of these clones. Taking this error into account, the sequence of one of the subgroup members (designated *BV20-1*) was found to be identical to the published sequence of BTB4. Two of the six sequences designated *BV20-3* were identical to BTB10, apart from one nucleotide difference leading to a coding substitution. In addition, the sequence designated *BV20-2* differed from the published sequence BTB1 by only five nucleotides. All but one (*BV20-4*) of the *BV20* sequences were represented in the cDNA clones derived from one of the animals in the study.

A feature of the *BV20* sequences was the presence of variability in the nucleotides that make up the three codons towards the C terminus, after the cysteine at position 93. In contrast to other subgroups, for which different clones of the same gene segment exhibited identical sequences in this region, variability was consistently observed between different *BV20* clones with otherwise identical sequences. Consequently, a consensus sequence could not be derived for the last three codons of the *V* region for most of the *BV20* subgroup members. These three codons were therefore excluded from the analyses of sequence differences between the *BV20* subgroup members discussed below.

Table 2 Comparison of sequence identities of members of the bovine *BV9* subgroup

	Number of clones	Nucleotide identities (%)											Human *BV9*
		9-2	*9-3*	*9-4*	*9-5*	*9-6*	*9-7*	*9-8*	*9-9*	*9-10*	*9-11*	*9-12*	
BV9-1	4	95.0	79.9	78.7	82.4	89.5	80.3	83.7	84.9	81.6	80.8	81.6	78
BV9-2	4		78.3	78.3	80.8	89.1	79.1	83.7	84.1	78.3	80.3	81.6	78
BV9-3	4			98.3	84.1	82.4	87.0	90.4	82.4	87.0	91.2	93.3	79
BV9-4	2				82.9	81.2	86.4	90.0	82.0	87.4	91.6	93.7	79
BV9-5	4					85.4	93.7	89.1	86.2	88.7	87.0	84.5	79
BV9-6	3						83.7	85.8	87.0	84.7	82.8	84.9	77
BV9-7	6							92.5	81.6	87.0	87.9	89.1	77
BV9-8	3								86.6	89.5	88.7	92.1	81
BV9-9	2									82.8	82.8	84.5	76
BV9-10	4										86.2	86.6	77
BV9-11	1											95.0	77
BV9-12	1												76

The human *BV9* (*BV1*) sequence used for comparison is contained in Concannon et al. (1986) [GenBank accession number M13836]. Sequences for *BV9-1*, *BV9-2*, *BV9-3*, *BV9-4*, *BV9-5*, *BV9-6*, *BV9-7*, *BV9-8* and *BV9-10* are also available on the GenBank database, with accession numbers AJ006568–76, respectively

Table 3 Comparison of sequence identities of members of the bovine *BV20* subgroup

	Number of clones	Nucleotide identities (%)							Human *BV20*
		20-2	*20-3*	*20-4*	*20-5*	*20-6*	*20-7*	*20-8*	
BV20-1	4	94.6	96.0	95.7	94.6	94.2	93.5	96.8	79
BV20-2	4		96.8	95.7	93.9	96.0	96.0	95.7	77
BV20-3	4			98.2	93.5	96.4	95.7	95.3	78
BV20-4	2				92.8	96.8	95.0	95.3	78
BV20-5	3					95.3	96.0	95.3	78
BV20-6	2						96.4	97.1	78
BV20-7	3							95.0	79
BV20-8	1								78

The human *BV20* (*BV2*) sequence human used for comparison is contained in Concannon et al. (1986) [GenBank accession number M13842]. Sequences for *BV20-1*, *BV20-3*, *BV20-5* and *BV20-7* are available on the GenBank database, with accession numbers AJ006581–83 and AJ006567, respectively

Sequence divergence within the *BV9* and *BV20* subgroups

The levels of identity in the nucleotide sequences between the members of the *BV9* and *BV20* subgroups are shown in Tables 2 and 3. The level of nucleotide identity between the *BV9* subgroup members (excluding *BV9-4*) ranged from 78.3 to 95.0%, with the majority of sequences showing 80–90% identity. The amino acid identity between these *BV9* subgroup members ranged from 64.2 to 91.0%. The *BV20* subgroup members exhibited higher levels of similarity than the members of the *BV9* subgroup. The percentage nucleotide identity between *BV20* subgroup members (excluding *BV20-4*) ranged from 92.8 to 97.1%, while the amino acid identity ranged from 84.0 to 94.6%.

Divergence in amino acid sequence within the *BV9* and *BV20* subgroups

Alignments of the predicted protein sequences of the *BV9* and *BV20* gene segments are illustrated in Fig. 1. Comparisons of the nucleotide sequences of the *BV9* subgroup members with *BV9-1* revealed that there was a preferential rate of coding vs non-coding substitutions in both the framework and predicted CDRs. The ratio of coding to non-coding substitutions ranged from 1.8:1 to 4.6:1 in the framework regions and from 3.5:1 to 9:1 in the CDR1 and CDR2 regions. In three of the sequences (*BV9-6*, *BV9-11* and *BV9-12*), all of the substitutions in CDR1 and CDR2 led to coding changes. The *BV20* subgroup members also showed a strong bias towards nucleotide substitutions that resulted in coding changes. In contrast to the findings with *BV9*, substitutions in

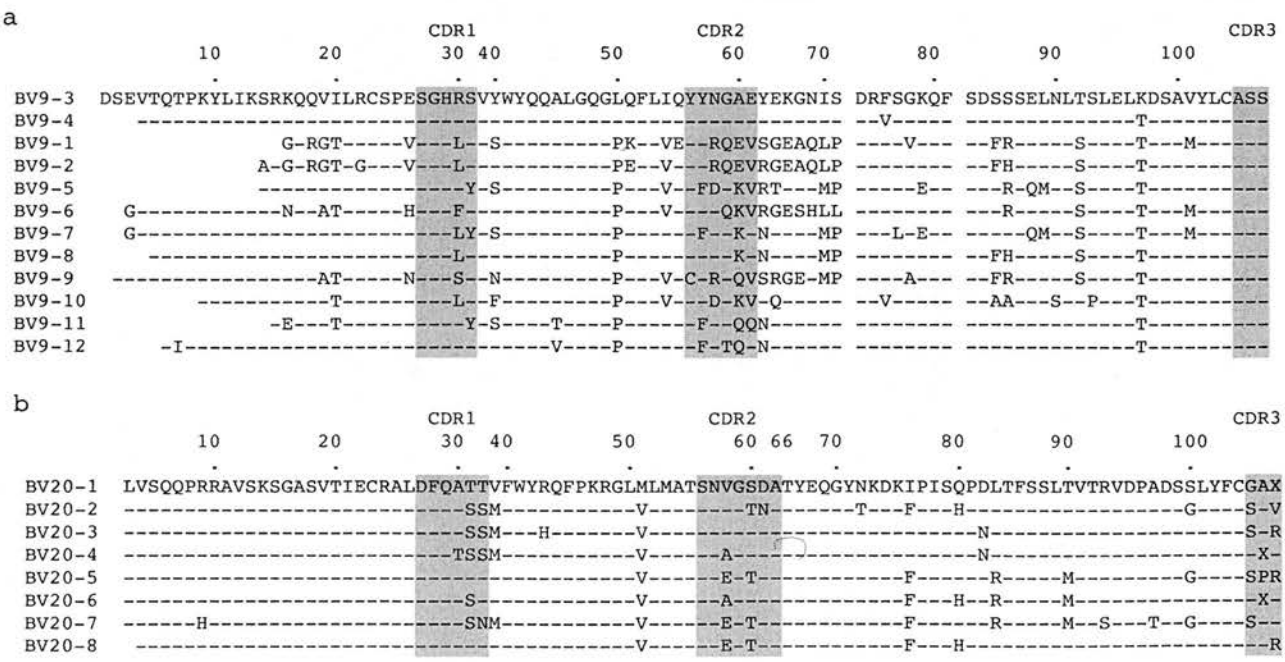


Fig. 1 Alignment of the amino acid sequences of members of the bovine *TCRBV9* (**a**) and *TCRBV20* (**b**) subfamilies. Identical residues are shown as *dashes*. Amino acid residues are numbered, and the CDRs are defined according to the IMGT system

both the framework and CDRs resulted in similarly high ratios of coding to non-coding changes, e.g. *BV20-6*—7:1 in the framework and 6:1 in the CDRs.

Demonstration of multiple gene segments in genomic DNA

To examine the complexity of the *BV9* and *BV20* subfamilies at the genomic level, Southern blot analysis was performed on non rearranged genomic DNA obtained from primary bovine fibroblast cell lines. In addition to probes specific for *BV9* and *BV20*, a *BV27* probe (representing a putative single-member subgroup) was used in these experiments. The four enzymes employed (*Hin*dIII, *Ssp*I, *Xba*I and *Sph*I) were selected on the basis that the sequenced *BV* gene segments belonging to these subgroups did not contain restriction sites for the enzymes. The results are shown in Fig. 2. Hybridisation with the *BV27* probe revealed a single band (5.0–7.9 kb in size) in preparations of genomic DNA digested with *Hin*dIII, *Ssp*I and *Sph*I. At the other extreme, the *BV9* probe gave a highly complex banding pattern, detecting at least 14 DNA fragments, ranging in size from 1.7 to 13.1 kb, in DNA digested with three of the four enzymes. The complexity of the banding pattern obtained with the *BV20* probe varied depending on the restriction enzyme employed. Thus, while only two prominent bands were detected in the *Hin*dIII-digested DNA, at least eight bands were discernible in the preparation digested with *Xba*I. There was a marked variation in the intensity of the bands detected with the *BV2* probe, suggesting that some of the bands may represent superimposition of more than one gene segment. Overall, the results are consistent with the presence of multiple *BV9* and *BV20* gene segments in bovine genomic DNA.

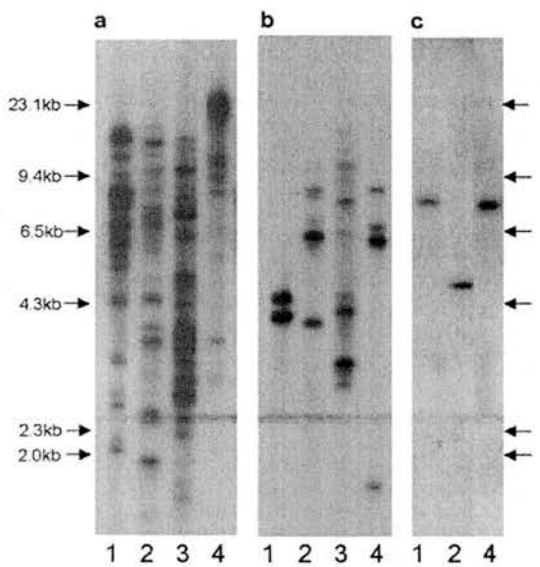


Fig. 2 Southern blot of bovine DNA digested with *Hin*dIII (*lane 1*), *Ssp*I (*lane 2*), *Xba*I (*lane 3*) and *Sph*I (*lane 4*), hybridised with ^{32}P-labelled probes for *BV9* (**a**), *BV20* (**b**) and *BV27* (**c**)

Discussion

The results presented, for the first time, in this paper provide an indication of the extent of the functional *TRBV* gene repertoire in cattle. In a previous study, Tanaka et al. (1990) identified members of nine bovine *TRBV* subgroups, but the small numbers of sequences did not permit a detailed classification. Thirty-nine distinct *TRBV* sequences were identified in the present study, including many that had previously been unreported, allowing an extended classification for bovine *TRBV* genes to be established. Sixteen *TRBV* subgroups were defined based on their sequence homology to each other and to human *TRBV* genes. These included the nine subgroups reported previously (Tanaka et al. 1990) and the seven new subgroups, *BV3*, *BV6*, *BV14*, *BV15*, *BV21*, *BV27* and *BV30*; the detailed sequences of which we have reported elsewhere (Houston and Morrison 1999). The *BV* gene sequences described in our study include subgroups in all of the phylogenetic lineages of *BV* genes described in humans and mice. Thus, our findings contradict proposals by Su et al. (1999) that cattle and sheep may have a more restricted repertoire of *BV* genes and that particular lineages of *BV* genes may not be expressed in these species.

Unlike rearranged immunoglobulin genes, there is no convincing evidence that TR variable genes undergo somatic mutation during maturation of immune responses. Thus, since the C-terminal codons, which are subject to modification during rearrangement of *BV* genes, were excluded from the analyses of the cDNA clones in this study, the 39 sequences identified are considered to represent distinct genomic *BV* gene segments. Two pairs of sequences (*BV9-3/BV9-4* and *BV20-3/BV20-4*) exhibited only four and five nucleotide differences, respectively. Based on experience with human *BV* genes, where sequences with five or less nucleotide differences have been shown to be alleles (Arden et al. 1995a,b), it is possible that these pairs represent alleles rather than distinct genes. However, in the absence of genetic data, their formal classification as alleles is not possible; they have therefore been tentatively designated as distinct subgroup members pending further studies. Together with the additional four sequences identified in previous studies (Table 1), this brings the total number of bovine *TRBV* gene segments so far identified to 41 (excluding the two potential allele sequences). Given the number of clones sequenced, this is unlikely to represent the full repertoire of functional *BV* genes in cattle. Nevertheless, the results indicate that the level of *TRBV* gene diversity in cattle is more similar to that in humans than in mice.

Our sequence analyses identified five bovine subgroups (*BV6*, *BV9*, *BV20*, *BV21* and *BV24*) that contained more than one member. Inclusion of previously published sequences (Buitkamp et al. 1993; Tanaka et al. 1990) revealed a further three subgroups (*BV7*, *BV19* and *BV29*) with more than one member. However, the most striking feature of our results was the large number of sequences in the *BV9* and *BV20* subgroups, both of which are single-member subgroups in man. Twenty of the 39 *BV* gene

sequences identified in the present study, including 61 (57%) of the cDNA clones examined, were classified within one of these two subgroups. The large number of *BV9* and *BV20* clones detected did not reflect over-representation of these segments in the activated T-cell populations since 10 of the 12 *BV9* sequences and 5 of the 8 *BV20* sequences (26 clones in total) were detected in the subset of 42 clones isolated from resting lymph node cells. That the different sequences represent distinct gene segments was supported by the finding that *BV9* and *BV20* probes hybridised with multiple bands in restriction enzyme digested genomic DNA. Variability in the complexity of the *BV20* banding patterns obtained with different enzymes is likely to be due to conservation of the flanking regions containing some of the restriction sites, reflecting the relatively recent duplication of these gene segments. Such a phenomenon has been observed for some mouse *TRAV* genes (Clark et al. 1995). This may also account for the finding by Tanaka et al. (1990) of a single *Pst*I DNA fragment in Southern blots of bovine DNA hybridised with *BV20* (*BV2* in former nomenclature).

Comparisons of *TRBV* sequences between species have provided convincing evidence that a set of ancestral *V* gene segments, recognisable as distinct *BV* gene subgroups, has been retained through speciation, forming the basis for the *TRBV* repertoire of each species. The presence of identifiable *BV* genes in birds (Tjoelker et al. 1990) and cartilaginous fish (Rast et al. 1997) indicates that the appearance of such genes predates evolution of mammals. Physical mapping studies of the *TRB* locus in mice and humans indicate that the relative order of the *TRBV* homologues is conserved although internal duplications have occurred in the human locus (Lai et al. 1988; Rowen et al. 1996). The results of the present study are consistent with this view of *TRBV* evolution, in that all of the sequences identified showed much greater sequence similarity to their human homologues than to other bovine *TRBV* subgroup members (data not shown). The level of sequence homology between the subgroups of different species varies, and identification of homologues may not always be straightforward. For example, the level of identity between mouse and human *TRBV* subgroups at the amino acid level ranges from greater than 70% for some subgroups to between 40 and 50% for others; this compares with identities below 40% at the amino acid level between different mouse *TRBV* subgroups (Arden et al. 1995b; Clark et al. 1995). All but one of the bovine *TRBV* subgroups showed greater than 60% identity to their human counterparts at the protein level, the exception being members of *BV21* (57%), while the average amino acid similarity between the genes in different bovine *BV* subgroups was less than 50%.

The number of *TRBV* gene segments encoded in the germline varies between species and includes both functional and non-functional genes. The functional repertoire is determined by both the number of subgroups that express functional genes and the number of functional genes within each subgroup. In humans, 32 subgroups have been defined, 23 of which contain a total of 47 functional gene segments (Arden et al. 1995a; Rowen et al. 1996), while in the mouse, 25 subgroups with 22 functional *TRBV* genes have been identified at the genomic level (Arden et al. 1995a). An additional 19 non-functional genes have been described in humans (Rowen et al. 1996). In contrast to mice, in which the majority of *TRBV* subgroups contain a single member, seven human *TRBV* subgroups are multimembered, the largest being *BV5* (five functional and two non-functional members), *BV6* (seven functional and two non-functional members) and *BV7* (six functional and three nonfunctional members) (Arden et al. 1995a; Rowen et al. 1996). In this respect, the bovine *TRBV* appears more similar to human than mouse, in that our results together with previously published data have identified eight multimember subgroups. However, the subgroups in which there has been amplification of gene segments differ between the two species; six of the eight bovine subgroups that contain more than one member are single-member subgroups in man, and two of these subgroups, *BV9* and *BV20*, have undergone extensive amplification in cattle. A recent study of more than 300 pig *BV* cDNA sequences identified 19 *BV* subgroups (Butler et al. 2005), which included all but one of the subgroups identified in cattle; clones within the subgroups in this study were not subjected to further classification. A more limited study in sheep identified 18 distinct *BV* sequences within 13 subgroups (Halsey et al. 1999), all but one of which (*BV12*, reported as *BV8* based on previous nomenclature) were identified in cattle. Interestingly, these genes include five distinct *BV9* sequences, suggesting that duplication of this subgroup may be a feature of ruminants.

Genomic sequencing of the human *TRBV* region has clearly demonstrated that multi-member subgroups of *BV* have arisen through gene duplication events (Rowen et al. 1996). The retention of duplications in particular *BV* subgroups indicates that they are likely to convey an evolutionary advantage for survival of the species. The α/β TR repertoire is generated in the thymus by production of T lymphocytes expressing different combinations of rearranged α- and β-chains that are subjected to sequential positive and negative selection processes based on the affinity with which they interact with different combinations of self MHC plus peptide (von Boehmer et al. 2003). Diversity is determined not only by the repertoire of *V* genes but also by various mechanisms that create diversity at the *V–D–J–C* junction; in the case of the β-chain, this involves the use of large numbers of *BJ* segments, imprecise joining of the segments, the cleavage and non-templated addition of nucleotides at the junctions of the gene segments (N region diversification) and the use of alternative reading frames for the *BD* genes, which combine to form the CDR3. Studies of the structure of the TR based on x-ray crystallography suggest that the CDR3 within the β-chain is principally involved in peptide interactions, whereas CDR1 and CDR2 predominantly form contacts with the α-helices bordering the peptide-binding groove of MHC molecules (Garboczi et al. 1996; Garcia et al. 1998; Kjer-Nielsen et al. 2003). Within a species, particular *BV* genes may either be of specific advantage because they give rise to TR specificities that allow effective responses to

prevalent life-threatening pathogens or convey more general benefits because of their ability to generate a diverse repertoire of TR specificities that survives negative selection in the thymus.

Acknowledgements The study was supported in part by a Ph.D. scholarship awarded by the Biotechnology and Biological Sciences Research Council to E.F. Houston. The authors also wish to thank Dr. Frank Katzer for assistance with the Southern blotting.

References

Akatsuka Y, Martin EG, Madonik A, Barsoukov AA, Hansen JA (1999) Rapid screening of T-cell receptor (TCR) variable gene usage by multiplex PCR: application for assessment of clonal composition. Tissue Antigens 53:122–134

Arden B, Clark SP, Kabelitz D, Mak TW (1995a) Human T-cell receptor variable gene segment families. Immunogenetics 42:455–500

Arden B, Clark SP, Kabelitz D, Mak TW (1995b) Mouse T-cell receptor variable gene segment families. Immunogenetics 42:501–530

Bernardin F, Doukhan L, Longone-Miller A, Champagne P, Sekaly R, Delwart E (2003) Estimate of the total number of $CD8^+$ clonal expansions in healthy adults using a new DNA heteroduplex-tracking assay for CDR3 repertoire analysis. J Immunol Methods 274:159–175

Bernardin F, Magierowska M, Dandekar S, Van Rompay KK, Delwart EL (2004) Number of $CD4^+$ and $CD8^+$ T-cell CDR3 clonotypes expanding during acute infection of macaques with simian immunodeficiency virus. Virology 322:105–117

Buitkamp J, Schwaiger, F-W, Epplen JT (1993) Vb6 T-cell receptor elements in artiodactyls: conservation and germline polymor phisms. Mammalian Genome 4:504–510

Butler JE, Wertz N, Sun J, Sacco RE (2005) Comparison of the expressed Vβ and Jβ repertoire of thymocytes and peripheral T cells. Immunology 114:184–193

Clark SP, Arden B, Kabelitz D, Mak TW (1995) Comparison of human and mouse T cell receptor variable gene segment subfamilies. Immunogenetics 42:531–540

Cohen GB, Islam SA, Noble MS, Lau C, Brander C, Altfeld MA, Rosenberg ES, Schmitz JE, Cameron TO, Kalams SA (2002) Clonotype tracking of TCR repertoires during chronic virus infections. Virology 304:474–484

Concannon P, Pickering LA, Kung P, Hood L (1986) Diversity and structure of human T-cell receptor β-chain variable region genes. Proc Natl Acad Sci U S A 83:6598–6602

Davis MM (1990) T cell receptor gene diversity and selection. Annu Rev Biochem 59:475–496

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programmes for VAX. Nucleic Acids Res 12:387–395

Garboczi DN, Ghosh P, Utz U, Fan QR, Biddison WE, Wiley DC (1996) Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. Nature 384:134–141

Garcia KC, Degano M, Pease LR, Huang M, Peterson PA, Teyton L, Wilson IA (1998) Structural basis of plasticity in T cell receptor recognition of a self peptide–MHC antigen. Science 279:1166–1172

Giudicelli V, Lefranc M-P (1999) Ontology for immunogenetics: the IMGT-ONTOLOGY. Bioinformatics 15:1047–1054

Halsey WA Jr, Palmer BE, DeMartini JC, Howell MD (1999) Analysis of sheep T cell receptor beta-chain heterogeneity. Immunogenetics 49:206–214

Houston EF, Morrison WI (1999) Identification of seven new *TCRBV* subfamilies in cattle (*Bos taurus*). Eur J Immunogenet 26:349–353

Ishiguro N, Tanaka A, Shinagawa M (1990) Sequence analysis of bovine T cell receptor β-chain. Immunogenetics 31:57–60

Kjer-Nielsen L, Clements CS, Purcell AW, Brooks AG, Whisstock JC, Burrows SR, McCluskey J, Rossjohn J (2003) A structural basis for the selection of dominant alphabeta T cell receptors in antiviral immunity. Immunity 18:53–64

Lai E, Concannon P, Hood L (1988) Conserved organization of the human and murine T-cell receptor beta-gene families. Nature 331:543–546

Lefranc M-P, Giudicelli V, Busin C, Bodmer J, Muller W, Bontrop R, Lemaitre M, Malik A, Chaume D (1998) IMGT, the international ImMunoGeneTics database. Nucleic Acids Res 26:297

Maniatis T (1992) Molecular cloning: a laboratory manual. Cold Spring Harbour Laboratory, Cold Spring Harbour

Morrison WI, McKeever DJ (1998) Immunology of infections with *Theileria parva* in cattle. In: Liew FW, Cox FEG (eds) Chemical immunology series, vol 70, Immunology of intracellular parasitism, Karger, Basel, pp 163–185

Rast JP, Anderson MK, Strong SJ, Luer C, Litman RT, Litman GW (1997) Alpha, beta, gamma, and delta T cell antigen receptor genes arose early in vertebrate phylogeny. Immunity 6:1–11

Rowen L, Koop BF, Hood L (1996) The complete 685-kilobase DNA sequence of the human β T cell receptor locus. Science 272:1755–1762

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74:5463–5467

Su C, Jakobsen I, Gu X, Nei M (1999) Diversity and evolution of T-cell receptor variable region genes in mammals and birds. Immunogenetics 50:301–308

Takeuchi N, Ishiguro N, Shinagawa M (1992) Molecular cloning and sequence analysis of bovine T-cell receptor gamma and delta chain genes. Immunogenetics 35:89–96

Tanaka A, Ishiguro N, Shinagawa M (1990) Sequence and diversity of bovine T-cell receptor β-chain genes. Immunogenetics 32:263–271

Tjoelker LW, Carlson LM, Lee K, Lahti J, McCormack WT, Leiden JM, Chen CL, Cooper MD, Thompson CB (1990) Evolutionary conservation of antigen recognition: the chicken T-cell receptor beta chain. Proc Natl Acad Sci U S A 87:7856–7860

Turner SJ, Diaz G, Cross R, Doherty PC (2003) Analysis of clonotype distribution and persistence for an influenza virus-specific $CD8^+$ T cell response. Immunity 18:549–559

von Boehmer H, Aifantis I, Gounari F, Azogui O, Haughn L, Apostolou I, Jaeckel E, Grassi F, Klein L (2003) Thymic selection revisited: how essential is it? Immunol Rev 191:62–78